SYNTACTIC VARIATION AND LINGUISTIC COMPETENCE: THE CASE OF AAVE COPULA ABSENCE

A DISSERTATION
SUBMITTED TO THE DEPARTMENT OF LINGUISTICS
AND THE COMMITTEE ON GRADUATE STUDIES
OF STANFORD UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Emily M. Bender
October 2000

I certify that I have read this dissertation and that in my opinion it is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Thomas Wasow
(Principal Co-Advisor)

I certify that I have read this dissertation and that in my opinion it is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Penelope Eckert
(Principal Co-Advisor)

I certify that I have read this dissertation and that in my opinion it is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

John Rickford

I certify that I have read this dissertation and that in my opinion it is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Ivan A. Sag

I certify that I have read this dissertation and that in my opinion it is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Arnold Zwicky

Approved for the University Committee on Graduate Studies:

Abstract

This thesis explores the implications for competence theories of syntax of the data on variation found by sociolinguists working in the Labovian tradition, through a case study of variable copula absence in African American Vernacular English (AAVE).

A distributional analysis of the categorical constraints on AAVE copula absence shows that it is indeed a syntactic, rather than phonological variable, contra Labov (1969, 1995). Further, its analysis requires a phonologically empty element, even the surface-oriented framework of Head-Driven Phrase Structure Grammar (HPSG) (Pollard and Sag 1994).

AAVE copula absence is also subject to well-studied and robust non-categorical grammatical constraints. Previous formal approaches to such non-categorical constraints on variation treat non-categorical grammatical constraints as separate from whatever social constraints might also apply. Building on the idea that variation is socially meaningful (Labov 1963, Eckert 2000), I propose that, on the contrary, social and grammatical constraints interact: social constraints are conceptualized as the social meaning of a variable, and grammatical constraints as the intensifying or attenuating effect of the grammatical environment on the social meaning or social value of the variable. This hypothesis is tested and substantiated by a matched-guise experiment, focusing on the effect of the following grammatical environment.

Three types of linguistic knowledge seem to be involved in the judgments the participants gave in the experimental task: knowledge of social meaning attached to linguistic forms, direct knowledge of a grammatical structure that is computable from more basic signs already in the grammar, and knowledge of the frequentistic,

non-categorical grammatical constraints on variation. Traditional conceptions of linguistic competence place all three of these types of knowledge outside the grammar proper. However, I argue that that distinction is not based on empirical evidence and should be subject to reevaluation. Further, I suggest that sign-based grammars are uniquely suited as models for exploring possible extensions of linguistic competence and that sociolinguistic variation, the social value of variables and the non-categorical grammatical constraints that apply to them provide an interesting locus for the study of the boundaries of linguistic competence.

Acknowledgements

First and foremost, I would like to thank my advisors Tom Wasow and Penny Eckert, and the rest of my committee, Ivan Sag, John Rickford, and Arnold Zwicky. A great privilege of being a student is having other people help develop one's ideas. With teachers like these, it was a privilege indeed. Their influence and guidance are apparent to me on every page of this thesis. They imparted their views of what language is, how it works, and how to study it, and I treasure this gift.

In addition, I would like to thank Tom for believing in me and my ideas from start to finish, for the way in which talking to him always brought out my best ideas, for being heroically prompt with feedback and for being a role model for how to lead an academic life. I would like to thank Penny for her enthusiasm for my topic, for always reminding me that things are never so simple, and for helping me to treat the social issues with subtlety. I would like to thank Ivan for welcoming me to Stanford and treating me as a colleague from the first, for reading my work critically, and for never being satisfied with pat answers. I would like to thank John for welcoming me to the copula lair and for sharing his vast expertise on this topic. Finally, I would like to thank Arnold for teaching a class on syntax and style just as I started this project, for sharing his joy at discovering this common interest, and for gently pointing out that the questions I asked could be fruitfully worked on for a lifetime, and, therefore, I couldn't expect to answer them all in my thesis.

I am grateful to the National Science Foundation for a Graduate Research Fellowship, to the Stanford Humanities and Sciences Graduate Alumni Association for a Dissertation Fellowship, and to the Department of Linguistics at Stanford for additional support. I am also grateful to the Department of Linguistics at Cornell

University for welcoming me as a visiting scholar during the final year of this project.

The experimental work reported in this dissertation would not have been possible without the help of many people. I am indebted to Jennifer Iljas and Mark Thomas for enthusiastically embracing the idea of research at a community college, and for getting the project approved, providing facilities to run the experiment and above all for allowing me and helping me to recruit students from Mark Thomas's class. I am grateful to Kristofer Jennings for his patient help with the statistics used in the analysis of the results. I came to him with some ideas of how to interpret the numbers. He turned those ideas into legitimate statistical tests and developed a computer program for running those tests, and then patiently explained to me how both worked. Finally, I would like to thank the many people who participated in my experiment, as speakers and listeners, for generously contributing their time, voices, and judgments, and for their interest in my project.

Likewise, the syntactic analysis would not have been possible without the patient cooperation of my four native-speaker consultants: Stacy Fambro, Andrea Kortenhoven, Sonja Lanehart, and Tracey Weldon. They tirelessly and promptly responded to seemingly endless sets of sentences arriving by email, not only giving their judgments but often offering other information about the sentences that guided my analysis and greatly sped up the process. I am also grateful to Stacy and Andrea for help early on as I designed the experiment. And finally, I thank Stacy for telling me her stories and introducing me to her friends, helping to me to feel less ignorant about the community whose language I have been studying.

My years at Stanford would not have been the joy that they were without the friends I made along the way: Susanne Riehemann, to whom I am grateful for much stimulating discussion on thesis-related topics and otherwise, and for the thrill of converging ideas; Tracy King, a role model and living proof that there is life after grad school; Dan Flickinger, who always brightens everyone's day and is a genuine pleasure to work with; Cathryn Donohue, Martina Faller, Andrea Kortenhoven, Eunjin Oh, Ida Toivonen, Jean Braithwaite, Jennifer Arnold, Yukiko Morimoto, Rob Malouf, Dave McKercher, Kathryn Campbell-Kibler, Mary Rose, Sarah Benor, Stacy Fambro, Chris Kuszmaul, and Raph Thomadsen.

To Vijay Menon, my husband of three months and soul mate of nine years, I am grateful for the million ways in which you have supported me throughout this project, starting with understanding when I chose to go to Stanford, 3,000 miles away from you. Whether across the country or right near by, you have helped me to keep my perspective and courage in the face of what seemed at times an impossible task. As we move on to new things and dream new dreams, you are my greatest source of happiness, my safe harbor, my home.

Finally, I would like to thank my parents Sheila Bender and Jim Bender, and my step-parents, Ann Venables and Kurt VanderSluis. Their love and support were not only instrumental in my finishing this project, but indeed made me who I am. Many years ago, they sent me to college saying, “Major in whatever you want. We’re sure you’ll make something of it.” Mom, Dad, Ann, Kurt: Here it is!

Contents

List of Tables

List of Figures

Chapter 1

Introduction

The original impetus for this project was the observation that generative grammar (especially generative syntax) seemed incompatible with the data on variation found by sociolinguists working in the Labovian tradition. It seemed to me that the intricate patterns of non-categorical grammatical constraints on variation might reflect some underlying linguistic knowledge. If they do, they pose several problems for generative theory, which is generally ill-equipped to deal with probabilistic or quantitative knowledge.

But how was I to show that the non-categorical grammatical constraints are indeed a matter of linguistic knowledge? They answer lay in the growing body of literature on the social value of variation (e.g., Labov 1963, Ochs 1992, Eckert 2000). I hypothesized that knowledge of non-categorical constraints would be apparent in their effect on the social value of sociolinguistic variables, and further that this effect would be detectable in a matched-guise experiment (Lambert et al. 1975).

The case study taken up here is variable copula absence in African American Vernacular English (AAVE). This variable was chosen because it is the single most well-studied syntactic variable and the non-categorical grammatical constraints are therefore well-established. This variable also turns out to have interesting syntactic properties.

The structure of this dissertation is as follows. Chapter 2 presents the theory of syntax assumed here (a version of HPSG, see Pollard and Sag 1994 and Ginzburg

and Sag to appear). Three aspects are highlighted. The first is that HPSG is a sign-based, surface oriented grammar. These design features are what makes it plausible to extend HPSG to the kinds of knowledge required to model non-categorical constraints on variation. The second aspect emphasized is the general lack of empty categories in HPSG (and related theories). It is against this background that the syntax of AAVE copula absence becomes particularly interesting. The third aspect highlighted is an account of the detailed properties of English auxiliaries, drawing on the extensive literature this topic has generated in HPSG and its precursors. This detailed background is required in order to develop and evaluate the possible analyses of copula absence discussed in Chapter 3.

Chapter 3 begins by establishing that copula absence is indeed a syntactic, rather than phonological, variable, contrary to the assertions of Labov (1969, 1995). Labov's claim is that AAVE simply carries the general English process of auxiliary contraction one step further, resulting in copula absence. For this to be true, contraction must feed deletion, and copula absence must therefore be possible only where contraction is. The data in Chapter 3 show that this is not the case; copula absence is possible in a couple of environments that exclude contraction. I then go on to consider four different syntactic analyses: two constructional analyses designed to avoid positing an empty category for AAVE copula absence, an analysis in terms of a phonologically empty verb, and an analysis in terms of a potentially phonologically empty construction. Somewhat surprisingly, given the general success of lexicalist frameworks at avoiding empty categories, the latter two analyses are shown to be the only adequate ones.

With this background on the syntax of AAVE copula absence, the next two chapters concentrate on the problem of non-categorical constraints on variation. Chapter 4 presents the matched-guise experiment. The results of this experiment provide preliminary evidence that AAVE speakers (and other African Americans who are familiar with, but do not speak, AAVE) have knowledge of the non-categorical constraint tested: the effect of the following grammatical environment.

Chapter 5 reviews previous approaches, formal and functional, to non-categorical constraints against the background of three aspects of the social value of variation: socially meaningful patterns of variation, socially meaningful individual tokens of variants, and the effect of the grammatical environment on the social value of variants observed in Chapter 4. The approaches considered are the Variable Rule approach of Labov (1969) and others, the Optimality Theoretic approach of Reynolds and Nagy (1994), Anttila (1997), and Boersma and Hayes (1999), and the functionalist approach of Kiparsky (1972, 1988). Of these, only the functionalist account is capable of modeling all three aspects of the social value of variation. However, in order to capture the results of the experiment, it must posit that speakers have knowledge of the (putatively functional) constraints behind the patterns of variation across different grammatical environments.

Chapter 6 discusses three types of linguistic knowledge that seem to be involved in the judgments the participants gave in the experimental task: knowledge of social meaning attached to linguistic forms, direct knowledge of grammatical structure that is computable from more basic signs already in the grammar, and knowledge of frequencies or probabilities. While all three are usually excluded from models of linguistic competence, there are in the literature proposals for including each in competence grammar, and sign-based grammars appear to be uniquely qualified to do so. The conclusion of this discussion is that the location of the boundaries of competence grammar should be considered an open issue.

Chapter 2

Background: Construction-HPSG

This chapter presents the theoretical background necessary to develop and evaluate the syntactic analyses of AAVE copula absence in Chapter 3. I will be assuming a theory that might be called 'Construction-HPSG', roughly Head-Driven Phrase Structure Grammar as presented in Pollard and Sag 1994, augmented with a theory of constructions, as in Sag 1997 and Ginzburg and Sag to appear, as well as the theory of lexical alternations developed in Koenig 1999. This theory is a particularly interesting one in which to compare the various analyses because of its surface-orientation.

Sections 2.1–2.2 give a brief introduction to HPSG. Section 2.3 reviews the analyses in HPSG and related frameworks that have shown that phonologically empty linguistic categories are not required for a range of linguistic phenomena. This section concludes with a brief overview of surface-orientation as a design feature of grammar. Section 2.4 introduces constructions, and Section 2.5 presents a theory of the properties of English auxiliaries.

2.1 The basics: types, constraints, and the core grammar

HPSG conceptualizes language as a system of signs, or pairings of (phonological and syntactic) form with (semantic and pragmatic) meaning. The grammar is merely a

static description of what signs are in the language, along with a basic mechanism for combining the signs.

More formally, an HPSG consists of a set of types, arranged into a multiple inheritance hierarchy, a set of constraints on those types, and an initial symbol. Each type is a (partial) description of a kind of linguistic object, such as a word or a phrase.[1] These types can be combined according the feature logic to create descriptions of larger linguistic objects. Any combination of types that satisfies the initial symbol is a description of a grammatical sentence of the language.

To make this concrete, let us consider a simple example. The analysis of (1) requires three word types (for *Kim, likes* and *bagels*) and two phrasal types (one to combine *likes* and *bagels* and one to combine *Kim* and *likes bagels*).

(1) Kim likes bagels.

Consider first the (partial) description of *likes* given in Figure 2.1.[2] This description is formalized as an attribute-value matrix (avm). The first attribute (or feature) is PHON(onology) and its value is a phonological representation of the sign. For ease of exposition, I will systematically use English orthography as a place holder for a more adequate phonological representation. Thus, the value of the feature PHON in Figure 2.1 is the English orthography 'likes'. The next feature is SYNSEM, specifying the syntactic and semantic properties of the word. Inside SYNSEM, the feature HEAD gives the part of speech (in this case, *verb*), and the further information that this verb is finite. The values of the next three features (SUBJ, SPR, and COMPS) specify that this verb is looking for one complement and a subject, but no specifier.[3] The boxed numbers (tags) in the values of SUBJ and COMPS indicate structure-sharing with the value of another feature, the argument structure (ARG-ST). The argument structure

[1] As we will see below, not all types are descriptions of linguistic units in this sense.

[2] This description and those that follow are partial in two respects: First, for ease of exposition, I am abstracting away from such issues as quantification, pied-piping, and extraction (although the latter will be addressed below) and therefore the avms shown lack certain features posited for those analyses. On quantification, see Copestake et al. 1999. On pied-piping, see Sag 1997 and Ginzburg and Sag to appear. Second, and more technically, the linguistic descriptions found in a grammar are usually underspecified to some extent with respect to the fully specified models that satisfy them. See Pollard and Sag 1994 for discussion.

[3] These are list-valued features, as indicated by the angle brackets. $\langle\ \rangle$ indicates the empty list.

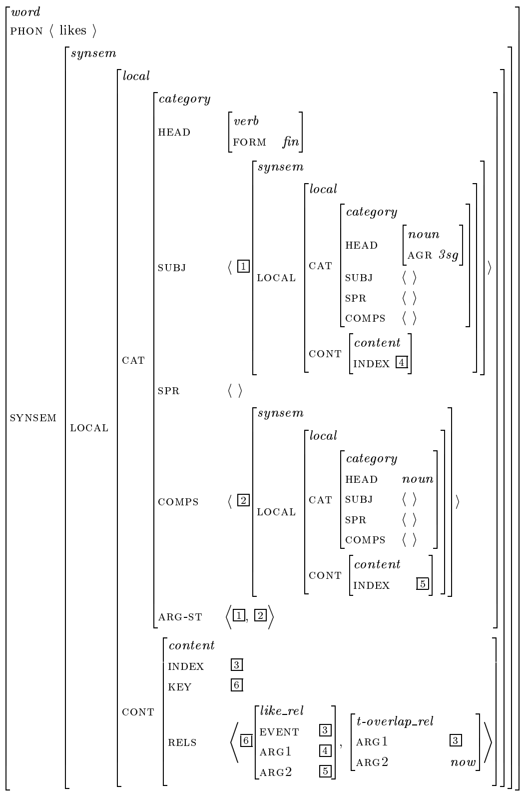


Figure 2.1: Lexical entry for *likes*

is a list of all of the dependents selected by this head, ordered in terms of obliqueness.[4] The substructure labelled $\boxed{1}$ (i.e., that which is also on the SUBJ list) specifies that the first syntactic argument of *likes* is a third singular noun phrase with the index $\boxed{4}$.[5] Here the term 'noun phrase' is shorthand for the specifications [HEAD *noun*] and [SUBJ ⟨ ⟩, SPR ⟨ ⟩, COMPS ⟨ ⟩]. That is, noun phrases are phrases headed by nouns with all of their valence requirements satisfied. Accordingly, the argument structure in Figure 2.1 is often abbreviated as in Figure 2.2.

Returning to the *synsem* of *likes*, the CONT(ent) value gives the semantics, in the style of Minimal Recursion Semantics (Copestake et al. 1999). Ignoring such issues as quantification, the semantics consists of two *relation*s, the *like_rel* and the *t(emporal)-overlap_rel*. The last of these is a representation of present tense,[6] stating that the situation described by the *like_rel* ($\boxed{3}$) overlaps with the present moment. The *like_rel* itself is a binary relation, whose two arguments ($\boxed{4}$ and $\boxed{5}$) are identified with the semantic indices of the two syntactic arguments. In the abbreviated notation shown in Figure 2.2, the tags corresponding to the semantic indices are given as subscripts to the arguments. The *content* value also has a feature KEY. This feature serves as a pointer to the main relation contributed by the word. Therefore its value is identified with one of the relations on the RELS list (here, $\boxed{6}$, the *like_rel*).

The label in italics at the top left of each pair of square brackets in Figure 2.1 gives the type of that substructure. The definition of each type includes which features are relevant to it (e.g., for the type *word*, the relevant features are PHON, SYNSEM and ARG-ST) as well as the type(s) of permissible values of each feature. For example, the value of SYNSEM is always of type *synsem* or one of its subtypes. When the value of a feature in a given description is the most general one possible for that feature, it is usually not indicated, as in Figure 2.2. As noted above, features can also take lists

[4] Argument structure is the locus of the binding theory, among other things. See Manning 1996 and Manning and Sag 1998 for the justification of Argument Structure as a level independent of both semantics and the valence features SUBJ/SPR/COMPS.

[5] Following Copestake et al. (1999), I am adopting a feature INDEX inside CONT. This feature is identified with the EVENT variable of the KEY for verbs and with the INST variable of the KEY for nouns. This feature is necessary because selecting heads need access to these variables in order to incorporate them properly into their own semantic relations.

[6] Others are possible.

as their values. Although this is not explicitly indicated in Figure 2.1, the lists are always lists of some particular type of object. For example, the value of the features SUBJ, SPR, COMPS and ARG-ST is always a list of *synsems*, and the value of RELS is always a list of *relations*.

$$\begin{bmatrix} \textit{word} & \\ \text{PHON} & \langle \text{likes} \rangle \\ \text{SYNSEM} & \begin{bmatrix} \text{LOCAL} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \textit{verb} & \\ \text{FORM} & \textit{fin} \end{bmatrix} \\ \text{SUBJ} & \langle \boxed{1} \rangle \\ \text{SPR} & \langle\, \rangle \\ \text{COMPS} & \langle \boxed{2} \rangle \\ \text{ARG-ST} & \left\langle \boxed{1}\text{NP}[\textit{3sg}]_{\boxed{4}}, \boxed{2}\text{NP}_{\boxed{5}} \right\rangle \end{bmatrix} \\ \text{CONT} & \begin{bmatrix} \text{INDEX} & \boxed{3} \\ \text{RELS} & \left\langle \begin{bmatrix} \textit{like\_rel} & \\ \text{EVENT} & \boxed{3} \\ \text{ARG1} & \boxed{4} \\ \text{ARG2} & \boxed{5} \end{bmatrix}, \begin{bmatrix} \textit{t-overlap\_rel} & \\ \text{ARG1} & \boxed{3} \\ \text{ARG2} & \textit{now} \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

Figure 2.2: Abbreviated lexical entry for *likes*

The lexical entries for *Kim* and *bagels* are given in Figures 2.3 and 2.4. Note that *Kim* has no valence requirements (SUBJ/SPR/COMPS are empty), and *bagels* has only an optional specifier (indicated by the parentheses around $\boxed{3}$ in its SPR value).

Now, the subcategorization information specified for *likes* is inert without phrase types to recognize it and pair it with nouns in the syntax. The first of these phrase types is the *head-complements-phrase*, shown in Figure 2.5. In HPSG, phrases, like words, are pairings of form and meaning. Accordingly, they have the features PHON and SYNSEM, just like words. Phrases differ from words, however, in that they encode constituent structure, via the HEAD-DTR and NON-HEAD-DTRS features.

$$\begin{bmatrix} \textit{word} & \\ \text{PHON} & \langle \text{Kim} \rangle \\ \text{SYNSEM} & \begin{bmatrix} \text{LOCAL} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \textit{noun} & \\ \text{AGR} & \textit{3sg} \end{bmatrix} \\ \text{SUBJ} & \langle\, \rangle \\ \text{SPR} & \langle\, \rangle \\ \text{COMPS} & \langle\, \rangle \\ \text{ARG-ST} & \langle\, \rangle \end{bmatrix} \\ \text{CONT} & \begin{bmatrix} \text{INDEX} & \boxed{1} \\ \text{KEY} & \boxed{2} \\ \text{RELS} & \left\langle \boxed{2}\begin{bmatrix} \textit{named\_rel} & \\ \text{INST} & \boxed{1} \\ \text{ARG} & \text{Kim} \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

Figure 2.3: Abbreviated lexical entry for *Kim*

$$\begin{bmatrix} \textit{word} & \\ \text{PHON} & \langle \text{bagels} \rangle \\ \text{SYNSEM} & \begin{bmatrix} \text{LOCAL} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \textit{noun} & \\ \text{AGR} & \textit{pl} \end{bmatrix} \\ \text{SUBJ} & \langle\, \rangle \\ \text{SPR} & \langle (\boxed{3}) \rangle \\ \text{COMPS} & \langle\, \rangle \\ \text{ARG-ST} & \langle (\boxed{3}\text{DetP}) \rangle \end{bmatrix} \\ \text{CONT} & \begin{bmatrix} \text{INDEX} & \boxed{1} \\ \text{KEY} & \boxed{2} \\ \text{RELS} & \left\langle \boxed{2}\begin{bmatrix} \textit{bagel\_rel} & \\ \text{INST} & \boxed{1} \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

Figure 2.4: Abbreviated lexical entry for *bagels*

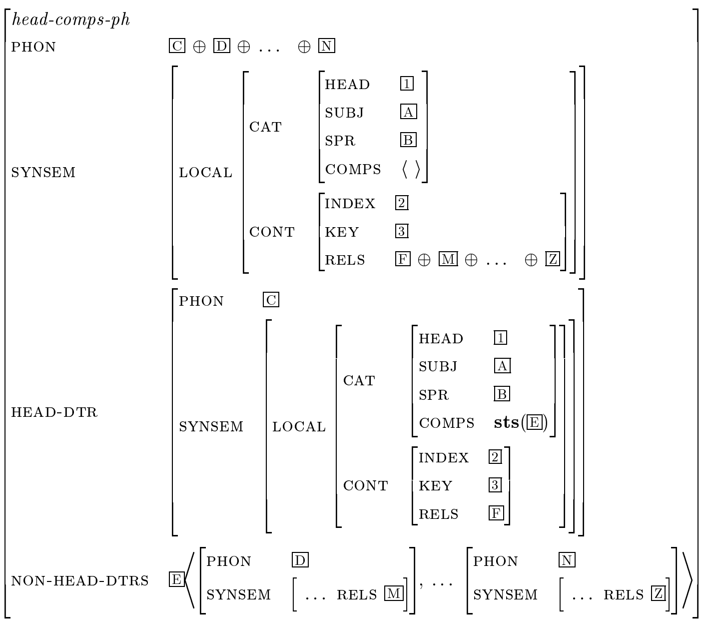


Figure 2.5: Head-complements phrase (abbreviated)

In Figure 2.5, the syntactic properties of the mother (that is, the value of SYNSEM | LOCAL | CAT) are systematically related to those of the head daughter. In particular, the HEAD, SUBJ and SPR values are shared between the two.[7] Only the COMPS value differs, as the mother is specified as [COMPS ⟨ ⟩]. This signifies that all of the complement requirements of the head have been realized within the phrase. In particular, they have been realized because each *synsem* that the head selects for (each *synsem* on its COMPS list) is identified with the SYNSEM value of one of the non-head daughters. (Figure 2.5 makes use of the 'sts' notation. The function **sts** (for 'signs-to-synsems') takes a list of *signs* as its argument and returns the corresponding list of *synsems*.) Whatever other valence requirements (SUBJ or SPR) there may be are not satisfied within this phrase, but 'passed up' to the next level. This is how subcategorization requirements are enforced.

The phonology and semantics of the phrase are constructed from the phonology and semantics of all of the daughters. In particular, the PHON value of the phrase is PHON value of the head daughter ($\boxed{C}$) followed by the PHON value of each of the non-head daughters ($\boxed{D}$ through $\boxed{N}$).[8,9] Similarly, the RELS value, or collection of semantic relations, of the mother is the append of the RELS value of the head daughter ($\boxed{F}$) to those of the non-head daughters ($\boxed{M}$ through $\boxed{Z}$). The INDEX and KEY values are identified with those of the head daughter.

Note that this formulation of the head-complements phrase is completely general: complement-taking words of any category can be the head daughter, and any number of selected complements can be accommodated. In the case of (1), the head daughter is the word *likes*, which selects for only one complement. Unifying the description of *likes* from Figure 2.1 with the value of HEAD-DTR in Figure 2.5 gives the description in Figure 2.6.

Since the description of *bagels* unifies with (that is, does not conflict with) the one non-head daughter called for by Figure 2.6, it can be combined with that structure

[7] Boxed tags with capital letters (e.g., $\boxed{A}$) are when the identified values are lists.

[8] Here and in what follows, the symbol $\oplus$ represents the 'append' relation:

$$\langle\, a, \ldots, b \,\rangle \oplus \langle\, c, \ldots, d \,\rangle = \langle\, a, \ldots, b, c, \ldots, d \,\rangle$$

[9] Other relationships between the PHON features are possible.

$$
\begin{bmatrix}
\textit{head-comps-ph} & \\
\text{PHON} & \boxed{C} \oplus \boxed{D} \\
\text{SYNSEM} & \begin{bmatrix} \text{LOCAL} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{HEAD} & \boxed{1} \\ \text{SUBJ} & \boxed{A} \\ \text{SPR} & \boxed{B} \\ \text{COMPS} & \langle\ \rangle \end{bmatrix} \\ \text{CONT} & \begin{bmatrix} \text{INDEX} & \boxed{2} \\ \text{KEY} & \boxed{3} \\ \text{RELS} & \boxed{E} \oplus \boxed{F} \end{bmatrix} \end{bmatrix} \end{bmatrix} \\
\text{HEAD-DTR} & \begin{bmatrix} \text{PHON} & \boxed{A}\langle\ \text{likes}\ \rangle \\ \text{SYNSEM} & \begin{bmatrix} \text{LOCAL} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{HEAD} & \boxed{1}\begin{bmatrix} \textit{verb} & \\ \text{FORM} & \textit{fin} \end{bmatrix} \\ \text{SUBJ} & \boxed{A}\langle\ \text{NP}[\textit{3sg}]_{\boxed{4}}\ \rangle \\ \text{SPR} & \boxed{B}\langle\ \rangle \\ \text{COMPS} & \langle\ \boxed{6}\text{NP}_{\boxed{5}}\ \rangle \end{bmatrix} \\ \text{CONT} & \begin{bmatrix} \text{INDEX} & \boxed{2} \\ \text{KEY} & \boxed{3} \\ \text{RELS} & \boxed{E}\left\langle \boxed{3}\begin{bmatrix} \textit{like\_rel} & \\ \text{EVENT} & \boxed{2} \\ \text{ARG1} & \boxed{4} \\ \text{ARG2} & \boxed{5} \end{bmatrix}, \begin{bmatrix} \textit{t-overlap\_rel} & \\ \text{ARG1} & \boxed{2} \\ \text{ARG2} & \textit{now} \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix} \\
\text{NON-HEAD-DTRS} & \left\langle \begin{bmatrix} \text{PHON} & \boxed{D} \\ \text{SYNSEM} & \boxed{6}\left[\text{LOCAL} \mid \text{CONT} \mid \text{RELS}\ \boxed{F}\right] \end{bmatrix} \right\rangle
\end{bmatrix}
$$

Figure 2.6: Head-complements phrase headed by *likes*

in a similar fashion. The PHON and SYNSEM values of the resulting description are shown in Figure 2.7.[10] (To keep this figure small enough to handle, the daughters features are suppressed.)

$$
\begin{bmatrix}
\textit{head-comps-ph} & \\
\text{PHON} & \langle\ \text{likes, bagels}\ \rangle \\
\text{SYNSEM} & \begin{bmatrix} \text{LOCAL} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \textit{verb} & \\ \text{FORM} & \textit{fin} \end{bmatrix} \\ \text{SUBJ} & \langle\ \text{NP}[\textit{3sg}]_{\boxed{4}}\ \rangle \\ \text{SPR} & \langle\ \rangle \\ \text{COMPS} & \langle\ \rangle \end{bmatrix} \\ \text{CONT} & \begin{bmatrix} \text{INDEX} & \boxed{2} \\ \text{KEY} & \boxed{3} \\ \text{RELS} & \left\langle \boxed{3}\begin{bmatrix} \textit{like\_rel} & \\ \text{EVENT} & \boxed{2} \\ \text{ARG1} & \boxed{4} \\ \text{ARG2} & \boxed{5} \end{bmatrix}, \begin{bmatrix} \textit{t-overlap\_rel} & \\ \text{ARG1} & \boxed{2} \\ \text{ARG2} & \textit{now} \end{bmatrix}, \begin{bmatrix} \textit{bagel\_rel} & \\ \text{INST} & \boxed{5} \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix} \end{bmatrix}
\end{bmatrix}
$$

Figure 2.7: Head-complements phrase *likes bagels*

The head-subject phrase, which combines *Kim* with *likes bagels*, is shown in Figure 2.8. As before, the syntactic features of the mother of this phrase (those under SYNSEM | LOCAL | CAT) are systematically related to the syntactic features of the head daughter. The HEAD, SPR and COMPS values are shared between the two. Further, the comps value of the head daughter must be empty. This ensures that

[10] In Figure 2.7, information that before was spelled out within the HEAD-DTR value is now spelled out within PHON and SYNSEM. Since the tags represent structure-sharing, this is only a notational change.

$$
\begin{bmatrix}
\textit{head-subj-ph} & \\
\text{PHON} & \boxed{B} \oplus \boxed{A} \\
\text{SYNSEM} & \begin{bmatrix} \text{LOCAL} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{HEAD} & \boxed{1} \\ \text{SUBJ} & \langle\, \rangle \\ \text{SPR} & \boxed{C} \\ \text{COMPS} & \boxed{D} \end{bmatrix} \\ \text{CONT} & \begin{bmatrix} \text{INDEX} & \boxed{2} \\ \text{KEY} & \boxed{3} \\ \text{RELS} & \boxed{E} \oplus \boxed{F} \end{bmatrix} \end{bmatrix} \end{bmatrix} \\
\text{HEAD-DTR} & \begin{bmatrix} \text{PHON} & \boxed{A} \\ \text{SYNSEM} & \begin{bmatrix} \text{LOCAL} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{HEAD} & \boxed{1}\begin{bmatrix} \textit{verb} & \\ \text{FORM} & \textit{fin} \end{bmatrix} \\ \text{SUBJ} & \langle \boxed{4} \rangle \\ \text{SPR} & \boxed{C} \\ \text{COMPS} & \boxed{D}\langle\, \rangle \end{bmatrix} \\ \text{CONT} & \begin{bmatrix} \text{INDEX} & \boxed{2} \\ \text{KEY} & \boxed{3} \\ \text{RELS} & \boxed{E} \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix} \\
\text{NON-HEAD-DTRS} & \left\langle \begin{bmatrix} \text{PHON} & \boxed{B} \\ \text{SYNSEM} & \boxed{4}\begin{bmatrix} \text{LOCAL} \mid \text{CONT} \mid \text{RELS} & \boxed{F} \end{bmatrix} \end{bmatrix} \right\rangle
\end{bmatrix}
$$

Figure 2.8: Head-subject phrase (abbreviated)

heads combine first with their complements, giving the constituent structure [*Kim* [*likes bagels*]] rather than [[*Kim likes*] *bagels*]. The SUBJ value of the mother is empty, as the subject is realized within the phrase. As before, the phrasal type enforces the selectional restrictions of the head: the SYNSEM value of the non-head daughter (the subject) is identified with the *synsem* on the head daughter's SUBJ list. This identification means that any requirements (such as agreement features) specified by the head daughter of its subject must be compatible with the subject. Similarly, the semantic index of the subject is available to the head daughter and (through a series of such identities, including $\boxed{3}$ in Figure 2.5) can be incorporated as an argument of a relation contributed by the head verb.[11]

In our analysis of (1), the head daughter of the head-subject phrase is the head-complements phrase *likes bagels*. Unifying the description of this phrase with the head daughter value of Figure 2.8 gives the description in Figure 2.9. (The HEAD-DTR and NON-HEAD-DTRS features of the head-complements phrase have again been left out for reasons of space.)

Since the description of *Kim* unifies with (that is, does not conflict with) the one non-head daughter called for by Figure 2.9, it can be combined with that structure in a similar fashion. The PHON and SYNSEM values of the resulting description are shown in Figure 2.10.

Figure 2.11 gives a 'translation' of the analysis of this sentence into the more familiar tree notation. In this figure, the nodes are labelled with the CAT values of the words and phrases postulated in the analysis. From this figure, it is apparent that constituent structure is minimal. Note also how the valence requirements of the verb get 'canceled off' as one moves up the tree.

[11] Unlike the head-complements phrase, this version of the head-subject phrase requires a finite verb as its head. As such, it corresponds to Sag's (1997) finite head-subject phrase. Sag notes that a complete grammar of English may require other types of head-subject phrase.

$$\begin{bmatrix} \textit{head-subj-ph} & \\ \text{PHON} & \boxed{B} \oplus \boxed{A} \\ \text{SYNSEM} & \begin{bmatrix} \text{LOCAL} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{HEAD} & \boxed{1} \\ \text{SUBJ} & \langle\ \rangle \\ \text{SPR} & \boxed{C} \\ \text{COMPS} & \boxed{D} \end{bmatrix} \\ \text{CONT} & \begin{bmatrix} \text{INDEX} & \boxed{2} \\ \text{KEY} & \boxed{3} \\ \text{RELS} & \boxed{E} \oplus \boxed{F} \end{bmatrix} \end{bmatrix} \end{bmatrix} \\ \text{HEAD-DTR} & \begin{bmatrix} \text{PHON} & \boxed{A}\langle \text{ likes, bagels } \rangle \\ \text{SYNSEM} & \begin{bmatrix} \text{LOCAL} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{HEAD} & \boxed{1}\begin{bmatrix} \textit{verb} & \\ \text{FORM} & \textit{fin} \end{bmatrix} \\ \text{SUBJ} & \langle\ \boxed{4}\text{NP}[\textit{3sg}]_{\boxed{5}}\ \rangle \\ \text{SPR} & \boxed{C} \\ \text{COMPS} & \boxed{D}\langle\ \rangle \end{bmatrix} \\ \text{CONT} & \begin{bmatrix} \text{INDEX} & \boxed{2} \\ \text{KEY} & \boxed{3} \\ \text{RELS} & \boxed{E}\left\langle \boxed{3}\begin{bmatrix} \textit{like\_rel} & \\ \text{EVENT} & \boxed{2} \\ \text{ARG1} & \boxed{5} \\ \text{ARG2} & \boxed{6} \end{bmatrix}, \begin{bmatrix} \textit{t-overlap\_rel} & \\ \text{ARG1} & \boxed{2} \\ \text{ARG2} & \textit{now} \end{bmatrix}, \begin{bmatrix} \textit{bagel\_rel} & \\ \text{INST} & \boxed{6} \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix} \\ \text{NON-HEAD-DTRS} & \left\langle \begin{bmatrix} \text{PHON} & \boxed{B} \\ \text{SYNSEM} & \boxed{4}\left[\text{LOCAL} \mid \text{CONT} \mid \text{RELS}\ \boxed{F}\right] \end{bmatrix} \right\rangle \end{bmatrix}$$

Figure 2.9: Head-subject phrase headed by *likes bagels*

$$\begin{bmatrix} \textit{head-subj-ph} & \\ \text{PHON} & \langle \text{ Kim, likes, bagels } \rangle \\ \text{SYNSEM} & \begin{bmatrix} \text{LOCAL} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \textit{verb} & \\ \text{FORM} & \textit{fin} \end{bmatrix} \\ \text{SUBJ} & \langle\ \rangle \\ \text{SPR} & \langle\ \rangle \\ \text{COMPS} & \langle\ \rangle \end{bmatrix} \\ \text{CONT} & \begin{bmatrix} \text{INDEX} & \boxed{2} \\ \text{KEY} & \boxed{3} \\ \text{RELS} & \left\langle \begin{bmatrix} \textit{named\_rel} & \\ \text{INST} & \boxed{5} \\ \text{ARG} & \text{Kim} \end{bmatrix}, \begin{bmatrix} \textit{like\_rel} & \\ \text{EVENT} & \boxed{2} \\ \text{ARG1} & \boxed{5} \\ \text{ARG2} & \boxed{6} \end{bmatrix}, \begin{bmatrix} \textit{t-overlap\_rel} & \\ \text{ARG1} & \boxed{2} \\ \text{ARG2} & \textit{now} \end{bmatrix}, \begin{bmatrix} \textit{bagel\_rel} & \\ \text{INST} & \boxed{6} \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

Figure 2.10: Head-subject phrase *Kim likes bagels*

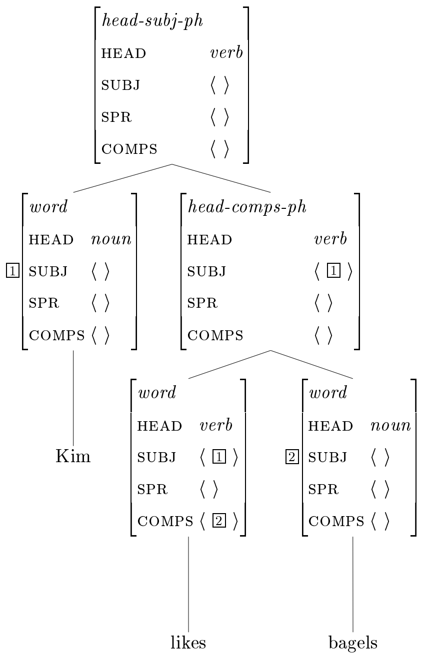


Figure 2.11: Tree version of *Kim likes bagels*

In this section, I have presented the analysis of this sentence in a bottom up fashion. It is important to note, however, that the partial descriptions of words and phrases used are defined in such a way that they may be combined in any order. That is, because this is a declarative, constraint-based grammar, one could get exactly the same result in any number of ways. For example, one could combine *Kim* with the NON-HEAD-DAUGHTER value of the *head-subj-ph*, then combine the *head-comps-ph* with the HEAD-DAUGHTER value of the *head-subj-ph*, then combine *bagels* with NON-HEAD-DAUGHTER value of the *head-comps-ph*, and plug *likes* in last as the head daughter of the *head-comps-phrase*. As long as all of the constraints provided by the partial description are consistent, they'll give the same result in any order. If any constraints conflict, it won't be possible to combine the types in any order.

2.2 Hierarchies

The types posited in the preceding discussion are quite detailed, as required by a precise formulation of the grammar. However, much of that detail is shared by different types, and one would like to capture such generalizations. In HPSG, this is accomplished by organizing the types into a hierarchy. For example, the *head-subj-ph* and the *head-comps-ph* are both subtypes of a more general type, *headed-phrase*. This mini type hierarchy is depicted as in Figure 2.12.

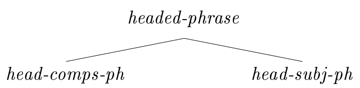


Figure 2.12: Miniature type hierarchy of headed phrases

The type *headed-phrase* bears all of the constraints that its subtypes have in common, as shown in Figure 2.13. In particular, *headed-phrase* specifies the identification of the HEAD and SPR values between mother and head daughter,[12] and the relationship of the PHON and CONT values of the mother to those of the daughters.[13] The subtypes inherit these constraints and add some of their own. This is illustrated for the *head-comps-ph* in Figure 2.14. Unifying all of the constraints given in Figure 2.13 with those given in Figure 2.14 gives exactly Figure 2.5 (modulo the specific numbers or letters chosen for tags, which are immaterial anyway).

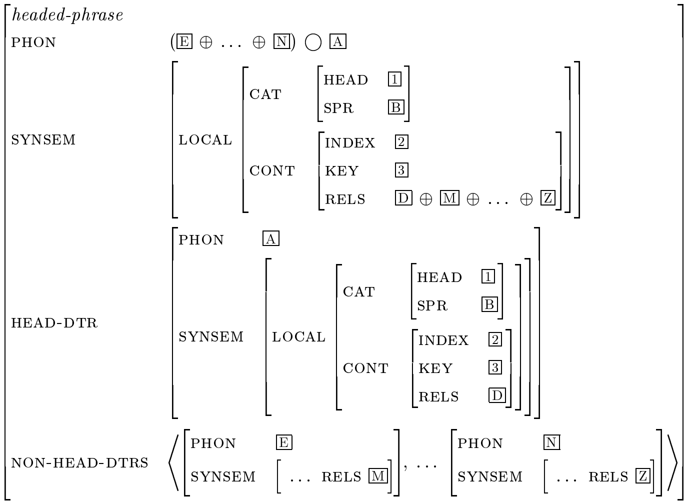


Figure 2.13: Headed phrase

[12]The relationship of the SUBJ and COMPS values of the mother to those of the head daughter can't be specified by *headed-phrase* as they differ for the two subtypes. Likewise, once a *head-spr-phrase* is included in the grammar, the same issue will arise for the SPR feature, and *headed-phrase* will not constrain the value of this feature. However, with a minimal use of defaults, Sag (1997) is able to state the general preservation of all valence features save those that are being canceled off as a constraint on *headed-phrase*.

[13]Here I've represented the PHON value of the mother as the list of PHON values of the non-head daughters with the PHON value of the head daughter shuffled in (◯) somewhere. The subtype *head-comps-ph* specifies that the head daughter's phonology comes first. A more general solution to this problem is to use linear precedence constraints, as in Pollard and Sag 1994.

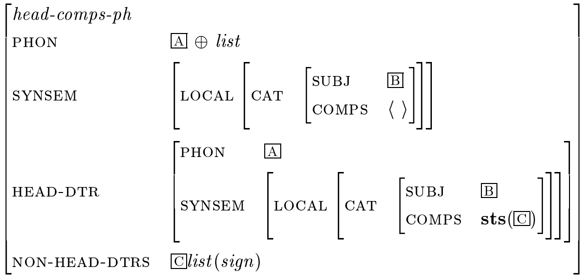


Figure 2.14: Head-complements phrase, as a subtype

Although the type *headed-phrase* is a part of the grammar, it is never used directly in the analysis of any sentence. In general, only 'maximal' word and phrase types (types with no subtypes) are so used.

In addition to organizing the types so that generalizations can be stated on supertypes, the hierarchy also plays a role in determining compatibility (i.e., whether two types can unify). For example, consider Sag and Wasow's (1999) sub-hierarchy of (Standard) English agreement types, given in Figure 2.15.[14]

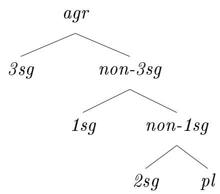


Figure 2.15: (Standard) English agreement types

These types are potential values of the feature AGR of nouns (cf. Figure 2.3,

[14]The features appropriate to these types are PERSON, NUMBER and GENDER. The first two are declared as appropriate for *agr*, and thus for all of its subtypes. In English, GENDER is only relevant for *3sg*.

page 10). Inflected verbs constrain the possible AGR values of their subjects. *Likes* will constrain its subject to be [AGR *3sg*] (cf. Figure 2.1). Since *3sg* and *non-3sg* are sister types (and have no common subtypes), *non-3sg* and any of its subtypes are incompatible with the *3sg* constraint imposed by the verb. On the other hand finite *like* requires that it's subject be [AGR *non-3sg*].[15] This is compatible with any subtype of *non-3sg*, so that an NP specified as, say, *1sg* could appear as the subject of *like*.[16]

Returning to the lexical types, much of the information given (especially for *likes* in Figure 2.1) represents generalizations that hold across all verbs or across subclasses of verbs. Here again, the generalizations are captured by factoring them out as constraints on supertypes. Beginning with the work of Flickinger et al. (1985), Flickinger (1987) and Pollard and Sag (1987), lexical types have been organized into a multiple inheritance hierarchy. Multiple inheritance hierarchies allow types to have more than one supertype and therefore express cross-cutting generalizations.

To take a simplified example, in English, valence possibilities cross-cut the part of speech distinctions. To represent this, part of the hierarchy for English could look like Figure 2.16. The subtypes of $\boxed{\text{VALENCE}}$ specify ARG-ST values (the feature name ARG-ST is omitted for reasons of space), while the subtypes of $\boxed{\text{POS}}$ specify HEAD values. (This is, of course, not an exhaustive listing of the subtypes in either dimension.) The next level down in the hierarchy specifies word classes, such as *verb-trans* which each inherit the constraints from one part of speech type and one valence type. Individual lexemes are subtypes of these word classes. Davis (1996) and Davis and Koenig (2000) extend this idea to a hierarchy of types that specify the linking between semantic arguments and the elements of ARG-ST. In a fully worked out grammar, all that needs to be stipulated for most lexical entries is the word class the entry belongs to, its phonological form, and the semantic relation it contributes.[17]

[15] Note that this is not true of the homophonous base form.

[16] For ease of exposition, I am representing agreement features as head features. In fact, agreement is much more complicated, involving both semantic agreement (mediated by indices) and syntactic agreement. For discussion, see Pollard and Sag 1994, Kathol 1999 and Bender and Flickinger 1999.

[17] The exceptions are words with idiosyncratic properties (defective paradigms, unusual complementation patterns, etc.). It would be superfluous to posit word classes in these cases each populated by only one lexical item. Instead, the lexical entries can inherit from more general word classes and specify the idiosyncratic information.

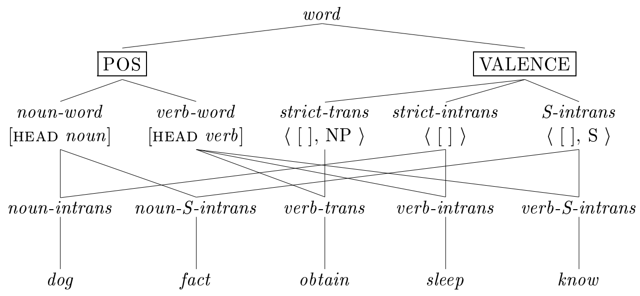


Figure 2.16: Multiple inheritance hierarchy

All of the rest of the complex information required (cf. Figure 2.1, page 7) is inherited from supertypes.

Note that it is rare in such multiple inheritance hierarchies to find all possible combinations of the supertypes instantiated. For example, in Figure 2.16, *strict-trans* has no mutual subtypes with *noun-word*. This is one sense in which the type hierarchy represents linguistically relevant *sub*generalizations.

Taking a completely underspecified type such as *object* as the root of the hierarchy, all of the types in the grammar, including lexical types, phrasal types and substructure types such as *synsem* or *3sg*, can be organized into one hierarchy. Figure 2.17 shows how this would look for the types introduced so far, plus a few contrasting types, such as *inf*(initival) which contrasts with *fin*(inite) under *form*. Because of space limitations, the subhierarchy under *word* is not included in this figure. Please refer to Figure 2.16 above.

This section has reviewed how types are organized into a hierarchy, and the information that that organization embodies. To summarize, the hierarchy allows for the expression of (possibly cross-cutting) generalizations and a specification of (in)compatibility between types.

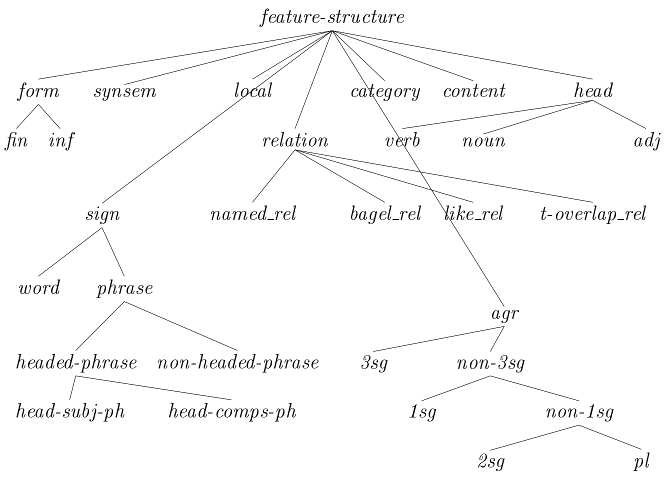


Figure 2.17: Hierarchy for types introduced so far

2.3 Grammar without empty categories

In this section, I will review lexicalist, constraint-based analyses of phenomena for which phonologically empty elements have been posited in other frameworks. In all three cases, it will be shown that empty categories are not required. These analyses are not original to HPSG, although I will give HPSG versions of them. The phenomena discussed are raising and control, pro-drop, and extraction/long-distance dependencies.

2.3.1 Raising and control

'Raising' and 'control' refer to the phenomena illustrated in (2) and (3), respectively:

(2) Kim tends to annoy Sandy.

(3) Kim tries to annoy Sandy.

In (2), *tends* provides a one-place relation, taking the proposition 'Kim annoys Sandy' as its sole argument. In (3), *tries* provides a two-place relation between Kim and the proposition 'Kim annoys Sandy'. In both cases, the subject of the higher verb is also interpreted as the subject of the lower verb. Transformational analyses of these phenomena (beginning with Rosenbaum 1967) posited underlying structures which included the sentence *Kim annoy Sandy.* The subject of the sentence is then either moved to be the subject of the higher verb (raising) or deleted under identity with the subject of the higher verb ('equi-NP-deletion'/control). More modern versions of these analyses (e.g., Chomsky 1981) still have movement in the raising sentence (leaving behind an NP-trace) and posit the phonologically empty element PRO as the embedded subject of the control sentence.

Bresnan (1978) suggests that instead the relationship of the matrix subject to the embedded verb could be mediated by the lexical entry of the raising or control verb (i.e., of *tries* or *tends*) itself. In HPSG, Pollard and Sag (1994) develop this analysis as shown in Figures 2.18 and 2.19, which give abbreviated lexical entries for *tend* and *try* respectively.[18]

$$\begin{bmatrix} \textit{lexeme} \\ \text{STEM } \langle \text{ tend } \rangle \\ \text{SYNSEM} \begin{bmatrix} \text{LOCAL} \begin{bmatrix} \text{CAT} \begin{bmatrix} \text{HEAD} & \textit{verb} \\ \text{SUBJ} & \langle \boxed{1}\text{NP} \rangle \\ \text{COMPS} & \left\langle \begin{bmatrix} \text{LOCAL} \begin{bmatrix} \text{CAT} \begin{bmatrix} \text{HEAD} & \begin{bmatrix}\textit{comp} & \\ \text{FORM} & \textit{inf}\end{bmatrix} \\ \text{SUBJ} & \langle \boxed{1} \rangle \\ \text{COMPS} & \langle\ \rangle \end{bmatrix} \\ \text{CONT} \begin{bmatrix}\text{INDEX} & \boxed{2}\end{bmatrix} \end{bmatrix} \end{bmatrix} \right\rangle \end{bmatrix} \\ \text{CONT} \begin{bmatrix} \text{INDEX} & \boxed{3} \\ \text{KEY} & \boxed{4} \\ \text{RELS} & \left\langle \boxed{4} \begin{bmatrix} \textit{tend\_rel} & \\ \text{EVENT} & \boxed{3} \\ \text{ARG1} & \boxed{2} \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

Figure 2.18: Raising: A lexical entry for *tend*

$$\begin{bmatrix} \textit{lexeme} \\ \text{STEM } \langle \text{ try } \rangle \\ \text{SYNSEM} \begin{bmatrix} \text{LOCAL} \begin{bmatrix} \text{CAT} \begin{bmatrix} \text{HEAD} & \textit{verb} \\ \text{SUBJ} & \langle \text{NP}_{\boxed{4}} \rangle \\ \text{COMPS} & \left\langle \begin{bmatrix} \text{LOCAL} \begin{bmatrix} \text{CAT} \begin{bmatrix} \text{HEAD} & \begin{bmatrix}\textit{comp} & \\ \text{FORM} & \textit{inf}\end{bmatrix} \\ \text{SUBJ} & \langle \begin{bmatrix}\text{INDEX} & \boxed{4}\end{bmatrix} \rangle \\ \text{COMPS} & \langle\ \rangle \end{bmatrix} \\ \text{CONT} \begin{bmatrix}\text{INDEX} & \boxed{2}\end{bmatrix} \end{bmatrix} \end{bmatrix} \right\rangle \end{bmatrix} \\ \text{CONT} \begin{bmatrix} \text{INDEX} & \boxed{3} \\ \text{KEY} & \boxed{5} \\ \text{RELS} & \left\langle \boxed{5} \begin{bmatrix} \textit{try\_rel} & \\ \text{EVENT} & \boxed{3} \\ \text{ARG1} & \boxed{4} \\ \text{ARG1} & \boxed{2} \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

Figure 2.19: Control: A lexical entry for *try*

Both *tend* and *try* select for a phrase headed by an infinitival complementizer (i.e., *to*), which has all of its complements saturated but which is still looking for a subject. The raising verb *tend* identifies its own subject ($\boxed{1}$) with the subject of its complement. The control verb *try* simply coindexes its own subject with the subject of its complement ($\boxed{4}$).[19] In either case, the index of the NP realized as the subject of the matrix verb gets passed along to the infinitival CP. In order to see how the index of *Kim* gets passed all the way to *annoy*, something needs to be said about the complementizer *to*. *To* itself is treated as a raising complementizer, with the lexical entry in Figure 2.20. As a semantically vacuous element, not only does *to* raise the subject of its complement, it also 'raises' or adopts the semantics, by identifying its INDEX value with that of its complement.

Using the familiar tree-structure notation, these lexical entries give rise to constituent structures like that shown in Figure 2.21. The lower clause needs no subject position, as the structure sharing in the lexical entries for the matrix verbs establishes all of the necessary identities. Since there is no subject position, there is no need for it to be filled with an NP-trace or PRO.[20]

[18]Because these types are descriptions of the uninflected lexemes, certain information that one would find in the descriptions of inflected words, such as tense information, is left out.

For ease of exposition, the syntactic arguments are specified in the SUBJ and COMPS values and the ARG-ST is suppressed. Since certain mismatches are allowed between ARG-ST and the valence features, lexeme descriptions usually don't stipulate anything about the value of SUBJ and COMPS. For more on this, see the discussion of pro-drop and extraction below.

[19]Thus there are two differences between raising and control verbs in HPSG: whether or not the subject plays a semantic role, and whether the subject is identified with, or merely coindexed with, the subject of the complement. This latter distinction is motivated by quirky case phenomena in Icelandic, where quirky case selected by the lower verb is only realized in the matrix clause if the matrix verb is a raising, rather than control, verb. On quirky case in Icelandic, see Andrews 1982 and Andrews 1990.

[20]Pollard (1989) posits a subtype of *synsem* called *PRO*, which appears on SUBJ and ARG-ST lists in environments where phrase-structure PRO appears as a subject on other accounts. The crucial difference is that *PRO* never occupies a phrase structural position. (See also Sag 1997:451.)

$$\begin{bmatrix} \textit{lexeme} & \\ \text{STEM} & \langle \text{ to } \rangle \\ \text{SYNSEM} & \begin{bmatrix} \text{LOCAL} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{HEAD} & \textit{comp} \\ \text{SUBJ} & \langle \boxed{1}\text{NP} \rangle \\ \text{COMPS} & \left\langle \begin{bmatrix} \text{LOCAL} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \textit{verb} & \\ \text{FORM} & \textit{base} \end{bmatrix} \\ \text{SUBJ} & \langle \boxed{1} \rangle \\ \text{COMPS} & \langle\, \rangle \end{bmatrix} \\ \text{CONT} & \begin{bmatrix} \text{INDEX} & \boxed{2} \\ \text{RELS} & \langle\, \rangle \end{bmatrix} \end{bmatrix} \end{bmatrix} \right\rangle \end{bmatrix} \\ \text{CONT} & \begin{bmatrix} \text{INDEX} & \boxed{2} \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

Figure 2.20: Raising: A lexical entry for *to*

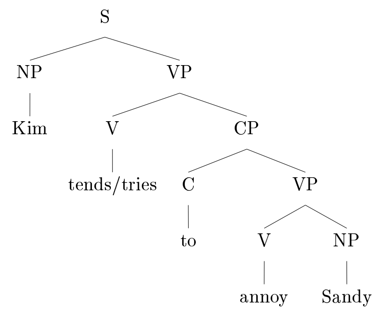


Figure 2.21: Constituent structure in raising/control sentences

2.3.2 Pro-drop

A second type of phonologically null element that has been posited is *pro*. This 'little pro', or phonologically empty pronoun, is posited whenever (NP) arguments are left unexpressed but aren't controlled by some other NP in the sentence (Chomsky 1981, Jaeggli 1982, Rizzi 1982). Classic examples are null subjects in Italian (4) and missing arguments in Japanese (5).

(4) Italian (Balari 1992:4)
Bisogna que parta subito
is.necessary that leave.1/3SG immediately
'It is necessary that I/he/she leave(s) immediately.'

(5) Japanese (Manning and Sag 1999:65)
Naoki-ga mi-ta.
Naoki-NOM see-PAST
'Naoki saw (it).'

However, Bresnan (1982) shows that one does not need to posit a phonologically empty pronoun (that occupies a position in the phrase structure) in order to account for this phenomenon. If a level of argument structure is available, it is sufficient to encode the pronominal properties there.

Thus, on the analysis sketched by Manning and Sag (1999), heads with 'missing' arguments have a mismatch between their ARG-ST value and the value of their valence features.[21] Thus *mita* 'saw' in (5) is subject to the constraints shown in Figure 2.22.

Such entries can be systematically produced either by lexical rule, or by underspecifying the relationship between ARG-ST and the valence features, on the analogy of Bouma et al.'s (in press) analysis of extraction (reviewed in §2.3.3 below). The constraint that relates the valence features to ARG-ST is called the Argument Realization Principle (ARP), and it can be formulated to allow for different kinds of mismatches. Figure 2.23 gives the ARP for a language like Japanese which allows any argument

[21] Balari's (1992) analysis predates the notion of argument structure in HPSG. Balari proposes instead that null subjects arise when an element appears on a SUBJ list, but is not realized in the phrase structure (or controlled by some other NP in the phrase structure).

$$\begin{bmatrix} \text{SYNSEM} & \begin{bmatrix} \text{LOCAL} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{SUBJ} & \langle \boxed{1} \rangle \\ \text{COMPS} & \langle\ \rangle \\ \text{ARG-ST} & \langle \boxed{1}\text{NP}[nom]_{\boxed{2}}, \text{NP}[pro]_{\boxed{3}} \rangle \end{bmatrix} \\ \text{CONT} & \begin{bmatrix} \text{KEY} & \begin{bmatrix} see\_rel & \\ \text{ARG1} & \boxed{2} \\ \text{ARG2} & \boxed{3} \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

Figure 2.22: Some constraints on *mita* with a missing object

to be missing. This constraint states that, for all signs of type *word*, the ARG-ST value must equal the SUBJ value followed by the COMPS value, with possibly some *pronoun_synsems* shuffled in. Figure 2.24 gives the ARP for a language like Italian which allows subjects to be missing.[22] This constraint states that, for all signs of type *word*, the ARG-ST value must equal the SUBJ value, plus a possibly empty list of *pronoun_synsems*, plus the COMPS value.[23,24]

$$\begin{bmatrix} word & \\ \text{SYNSEM} \mid \text{LOCAL} \mid \text{CAT} & \begin{bmatrix} \text{SUBJ} & \boxed{A} \\ \text{COMPS} & \boxed{B} \\ \text{ARG-ST} & \boxed{A} \oplus \boxed{B} \bigcirc list(pron\_synsem) \end{bmatrix} \end{bmatrix}$$

Figure 2.23: Argument Realization Principle for Japanese-type language

[22] As Balari (1992) argues at length, the picture is somewhat messier: even Italian has some subjects (e.g., second person singular subjects in embedded subjunctives) which cannot be dropped. However, I see no reason that an Argument Realization analysis could not be scaled up to include for such constraints.

[23] $\ominus$ indicates the relation of list subtraction. In the following statement, Z is a list that results from removing some number (possibly zero) of elements from X while preserving the order of the elements with respect to each other. Y is the list of elements removed.

$$X \ominus Y = Z$$

It follows that in Figure 2.24, the first part of the argument structure ($\boxed{A}$) is the SUBJ list plus some number of *pronoun_synsems*. The list $\boxed{A}$ must also be constrained to contain at most one element.

[24] Both of these ARPs will need to be revised to allow for extraction as well as pro-drop. See note 31 on page 38.

$$\begin{bmatrix} \textit{word} \\ \text{SYNSEM} \mid \text{LOCAL} \mid \text{CAT} \begin{bmatrix} \text{SUBJ} & \boxed{A} \ominus \textit{list}(\textit{pron\_synsem}) \\ \text{COMPS} & \boxed{B} \\ \text{ARG-ST} & \boxed{A} \oplus \boxed{B} \end{bmatrix} \end{bmatrix}$$

Figure 2.24: Argument Realization Principle for Italian-type language

These two examples (raising/control and pro-drop) show that the information encoded by argument structure is akin to that encoded by the deep structure in early versions of transformational grammar. Argument structure encodes the possible syntactic arguments of a head and their 'underlying' obliqueness relationships. There are different possible relationships between argument structure and the surface configurations the head can appear in.[25] The advantage of a theory with a level like argument structure (or f-structure in LFG) is that the similarities between the different realizations of the arguments of the same head can be described without positing phonologically null syntactic elements.

2.3.3 Extraction

Since Wasow 1972 (see also Lasnik 1992), transformational analyses of extraction phenomena (topicalization, fronted *wh*-questions, *tough*-movement) have posited an empty category, called a trace, at the bottom of the extraction dependency. For example, the movement that creates the extraction dependency in (6) relates the underlying structure (a) to the surface structure (b) in Figure 2.25.[26] The matching subscripts on *bagels* and the trace in the surface structure indicate that *bagels* binds the trace.

(6) Bagels I know Kim likes.

Traceless analyses of extraction have been proposed in various frameworks by Ades and Steedman (1982) (Categorial Grammar), Gazdar et al. (1984) (GPSG), Kaplan

[25] See Manning and Sag 1998 for a discussion of the range of possible relationships between ARG-ST and the valence lists.

[26] Since this discussion is focussed on traces, Figure 2.25 abstracts away from the CP/IP notation and the associated empty categories.

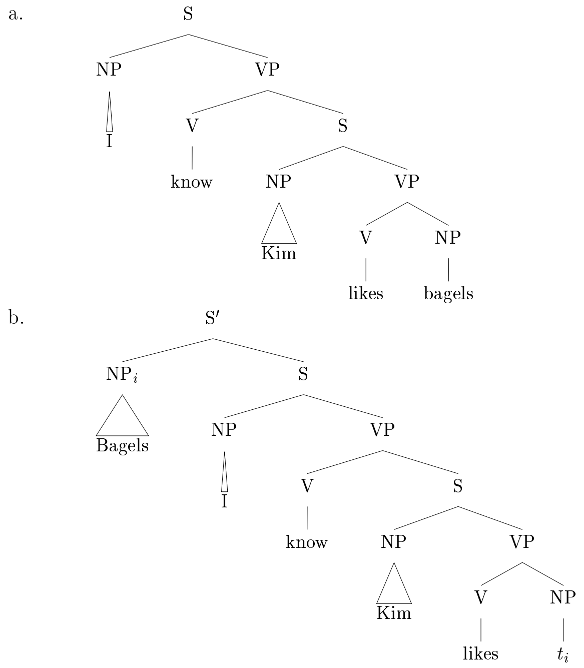


Figure 2.25: Movement analysis of extraction

and Zaenen (1988) (LFG), and Pollard and Sag (1994) (HPSG), among others. The history of the GPSG/HPSG version of traceless extraction proceeded in two steps. The first step was Gazdar's (1981) demonstration that a context-free grammar with complex category labels could model extraction, without recourse to transformations. The essence of Gazdar's analysis is that nodes dominating a trace carry that information 'up the tree' to the subtree that combines with the filler of the trace. That is, a sentence with an NP gap in it is a different category (S/NP or 'S slash NP') from a sentence with no gap (just plain S), and there is a special phrase structure rule that combines a filler NP with an S/NP. Under this analysis, the single structure representing (6) is as in Figure 2.26.[27]

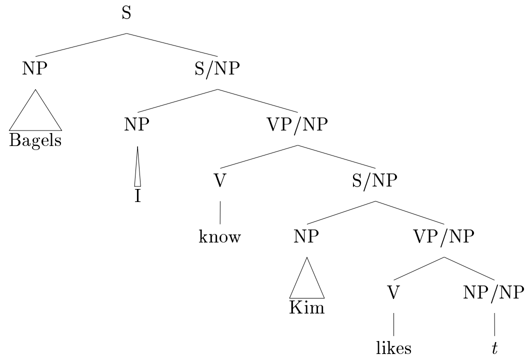


Figure 2.26: Slash analysis of extraction

When viewed in this light, extraction constructions involve three separate parts: the bottom, where something is missing, the middle, where that information is transmitted, and the top, where the filler is found (Gazdar et al. 1985:138). Since there is no movement on this analysis, it is possible to consider other formulations of the

[27] *Bagels* and the trace (*t*) are not coindexed this time, as "*t* is a dummy element [which] serves no semantic function ... " (Gazdar 1981:162) Gazdar's system includes separate semantic rules associated with each phrase structure rule which conspire to link up the interpretation of *bagels* with an argument of *likes*' relation.

bottom of the dependency which do not involve a trace or even a phrase structural position that a trace could occupy. The various proposals to this effect constitute step two.

Pollard and Sag's (1994), Ch. 9 proposal (refined by Sag (1997) and Bouma et al. (in press)) makes use of the potential for mismatch between the argument structure and the valence features. In particular, they posit word types such as the one in Figure 2.27.[28]

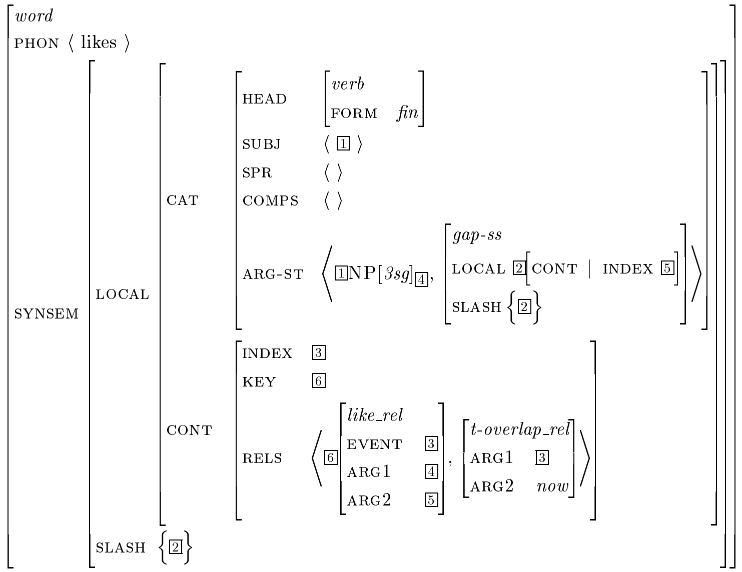


Figure 2.27: Lexical entry for *likes* with an extracted complement

[28] This is an example of complement extraction. The same mechanism can be extended to adjunct extraction (see Pollard and Sag 1994:384–388 and Bouma et al. in press). For a discussion of traceless subject extraction, see §3.4.3 below.

The lexical item in Figure 2.27 is not seeking any complements (its COMPS list is empty), but it does have a non-empty SLASH value:[29] The LOCAL value of its second argument is entered into the SLASH set of this argument (and by Sag's (1997) lexical amalgamation version of slash-passing) into the SLASH set of the verb. This information will be transmitted up the tree and eventually identified with the LOCAL value of the filler. Since the LOCAL value contains both the CAT and CONT substructures, any syntactic or semantic requirements the verb places on its second argument will also be transmitted.

Translating again to the more familiar tree notation, the constituent structure on this analysis is as in Figure 2.28. This constituent structure involves no trace and no phrase-structural position for the complement of *likes*.

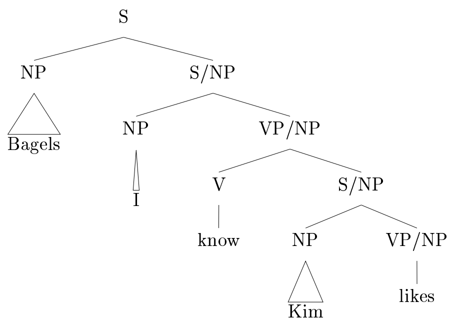


Figure 2.28: Traceless analysis of extraction

Pollard and Sag (1994) related lexical entries such as in Figure 2.27 to more basic lexical entries by means of a lexical rule. Bouma et al. (in press) model the relationship between the two types of lexical entries by instead underspecifying the relationship between ARG-ST and the valence features. In particular, they posit a

[29]The feature SLASH shown in Figure 2.27 is feature-based encoding of Gazdar's complex symbols with '/'.

constraint on the type *word* such as in Figure 2.29.[30] This constraint allows for a mismatch between the COMPS list and the non-initial members of the ARG-ST list. In particular, it states the the COMPS list may be shorter than the corresponding part of the ARG-ST list ($\boxed{2}$), just in case the elements 'removed' are unified with the type *gap-ss*. The type *gap-ss* (a subtype of *synsem*) identifies its LOCAL value with the sole element of its SLASH set, as in Figure 2.30, starting off the slash dependency.

$$\begin{bmatrix} \textit{word} \\ \text{SYNSEM} \begin{bmatrix} \text{LOCAL} \begin{bmatrix} \text{CAT} \begin{bmatrix} \text{SUBJ} & \boxed{A} \\ \text{COMPS} & \boxed{B} \ominus \textit{list(gap-ss)} \\ \text{ARG-ST} & \boxed{A} \oplus \boxed{B} \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

Figure 2.29: Argument Realization Principle

$$\begin{bmatrix} \textit{gap-ss} & \\ \text{LOCAL} & \boxed{1} \\ \text{SLASH} & \left\{ \boxed{1} \right\} \end{bmatrix}$$

Figure 2.30: *gap-ss*

Both the descriptions of *likes* in Figure 2.2 (page 9) and in Figure 2.27 satisfy the constraint in Figure 2.29, so both are licensed.[31]

[30]This presentation follows Bouma et al. (in press) in abstracting away from the feature SPR for simplicity, but differs in that their feature DEPS is replaced by ARG-ST. The value of DEPS is the append of the value of ARG-ST plus a (possibly empty) list of adjuncts. This is crucial for Bouma et al.'s (in press) uniform account of adjunct and complement extraction.

Note also that the constraint on (English) verbs that their SUBJ lists be at most one element long is stated elsewhere in the type hierarchy.

[31]The Argument Realization Principle for a null complement language such as Japanese would have to also include the constraint in Figure 2.23, page 32. The result would look like this:

$$\begin{bmatrix} \textit{word} \\ \text{SYNSEM} \begin{bmatrix} \text{LOCAL} \begin{bmatrix} \text{CAT} \begin{bmatrix} \text{SUBJ} & \boxed{A} \\ \text{COMPS} & \boxed{B} \ominus \textit{list(gap-ss)} \\ \text{ARG-ST} & \boxed{A} \oplus \boxed{B} \bigcirc \textit{list(pron\_synsem)} \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

2.3.4 Summary

This section has reviewed lexicalist, constraint-based analyses which show that empty categories are not required in the description of raising/control, pro-drop, or even extraction. In the last case especially, there is both linguistic and psycholinguistic evidence to prefer the traceless analysis.[32]

The most striking piece of linguistic evidence, as noted by Sag (1998), is the Conjunct Constraint subclause of Ross's (1967)b Coordinate Structure Constraint:

(7) In a coordinate structure, no conjunct may be moved.

This is illustrated by the examples in (8), from Sag 1998.

(8) a. *Which of her books did you read both [[a review of __] and [__]]?
b. *Which of her books did you find [[__] and [a review of __]]?
c. *Which rock legend would it be ridiculous to compare [[__] and [__]]? (cf. Which rock legend would it be ridiculous to compare __ with himself?)

Sag notes that these facts follow directly from a traceless analysis of extraction:

> A *Wh*-gap is simply a position where an element selected by a head (whether complement or adjunct) fails to be realized [. . .].
>
> The elements that are coordinated, i.e. the conjuncts of a coordinate structure, must be syntactic constituents (or perhaps sequences thereof).
>
> Conjunctions are not heads; rather, coordinate structures instantiate an independent construction type.
>
> Therefore, *wh*-gaps, which are not constituents, can never be conjuncts. (Sag 1998:5)

The psycholinguistic evidence is due to Pickering and Barry (1991) who note the difference in difficulty of processing between (9) and (10).

(9) That's the prize which we gave [every student capable of answering every single tricky question on the details of the new and extremely complicated theory about the causes of political instability in small nations with a history of military rulers] __ .

[32]For arguments against putative evidence for traces, see Sag and Fodor 1994 and Sag 1998.

(10) We gave [every student capable of answering every single tricky question on the details of the new and extremely complicated theory about the causes of political instability in small nations with a history of military rulers] [a prize].

Pickering and Barry (1991) argue that if the interpretation of the extracted element in (9) is processed only once the (putative) trace is encountered, it should be as difficult to process as (10). They take the fact that (9) is relatively easy to process as evidence that the interpretation of the extracted element is processed at the selecting head (*gave*), i.e., evidence against a trace-based theory of extraction.

More generally, not only traces, but all empty elements run counter to the surface-orientation of lexicalist, constraint-based theories of grammar. According to Sag and Wasow (1999), a grammar is surface-oriented if it "provides a reasonably simple structure that is directly associated with the string of words that constitute each sentence" (Sag and Wasow 1999:219). Surface-orientation is one of the design features of what they call 'performance-plausible competence grammar'. It is motivated by evidence that language processing is incremental, with hearers producing partial structures and partial interpretations of an incoming speech stream syllable by syllable. In order for a model of competence grammar to be able to be embedded in a performance system that can model this kind of incremental processing, the less abstract the syntactic representations, and the more closely said representations are tied to the surface string, the better. A parser that had empty categories in its lexicon must allow for their appearance at every point in the surface string. Surely this would interrupt incremental processing. At the very least, it would be inefficient.[33]

2.4 Constructions

Early work in HPSG (e.g., Flickinger et al. 1985, Flickinger 1987, Pollard and Sag 1987 and Pollard and Sag 1994) emphasized moving grammatical information into complex

[33]Johnson and Kay (1994) show that a parser can terminate even given empty categories as possible elements of the input string as long as all empty categories are sponsored by some other lexical item. That is, the lexical entries for the pronounced lexical items would specify how many empty categories each pronounced item could sponsor, bounding the number of empty categories that need to be considered. Guaranteed termination and efficiency are two different issues, however.

lexical entries (arranged in a hierarchy) that interacted with maximally simple phrase structure schemata. These schemata (which numbered only six in Pollard and Sag 1994) projected the information borne by the lexical entries into the rich variety of sentences found in language.

Although this strategy had impressive initial success, later work in HPSG (e.g., Sag 1997 and Ginzburg and Sag to appear) as well as work in Construction Grammar (e.g., Fillmore 1987, Fillmore 1988 and Fillmore et al. 1988; see also Zwicky 1994) has shown that as one moves away from the 'core' aspects of grammar, more complex phrase structure schemata (constructions) are required. In fact, this move is completely compatible with the notion of sign based grammar, as developed in Pollard and Sag 1994: If words and phrases are both just types of signs, and signs are pairing of form and meaning, then there is no reason for phrasal signs not to include more interesting pairings of form and meaning than the bare, simple phrase structure schemata assumed before. Further, as shown by Sag (1997), by extending the idea of the multiple inheritance hierarchy from words to phrases, cross-constructional generalizations can be captured in a theory with multiple, specific phrase structure types. In this section, I will briefly review two examples of the use of constructions in linguistic analysis: The analyses of the correlative conditional construction in Fillmore (1987, 1988) and McCawley (1988), and Sag's (1997) constructional analysis of relative clauses in English.

2.4.1 Correlative conditionals

The English correlative conditional construction, illustrated in (11–14), provides one of the most compelling arguments for the need for a notion of construction in grammatical theory. This construction is peripheral in the sense that its intricacies go beyond what can be captured by most models of "core" grammar, but far from peripheral in that it does not seem to be restricted to rarely used registers. The examples in (11) and (12) come from the Callhome corpus of spoken American English and the examples in (13) and (14) come from the North American News Text corpus, and

in particular from the New York Times.[34] Further, this construction is not some peculiarity of English alone. As McCawley (1988) argues, German and Mandarin have constructions that fulfill similar functions and similar, although not identical, constraints.

(11) Anyway, the more he hears about the community, the more he hears it's not so great.

(12) And also I think the sunnier, the more bright it is, the fuzzier it comes out for some reason.

(13) It was a fragile hope that began as an impossibility and gradually assumed a reality all its own; the further she marched through the draw, the more it seemed that a 10th championship was not only her desire but her destiny.

(14) He said the study found that most accused doctors were between ages 40 and 69. "The older the physician, the more likely the resolution would be surrender of the license," Winn wrote.

Fillmore (1987, 1988) and McCawley (1988), discuss numerous properties of this construction, including the following:[35]

- It has two daughters.
- Each daughter is a filler-gap construction, with a left-dislocated filler.
- The filler of each daughter is the word *the* followed by a comparative phrase.
- The antecedent is optionally a negative polarity context (15).

(15) a. The more you do any of that, the easier it gets. (Fillmore 1987:165)
b. The more noises I hear, the more sure I am that someone is here.
c. #The more noises I hear, the more sure I am that anyone is here.

[34] Both of these corpora are available from the Linguistic Data Consortium: http://www.ldc.upenn.edu.

[35] This presentation is restricted to the non-inverted (antecedent first) variety of correlative conditional, for simplicity.

- The *will* future is disallowed in the antecedent (16).

(16) The faster you (*will) drive, the sooner you'll get there.
(McCawley 1988:177)

- The copula can be omitted, just in case its complement is the extracted element and the correlative conditional is of the generic type of correlative conditional (17).

(17) a. The more outrageous a politician's promises (are), the bigger his vote count (is).
b. The more obnoxious Fred *(is), the less attention you should pay to him.
c. The happier the customers *(are) behaving, the more things you should try to sell them. (McCawley 1988:178)

Some of these properties follow from the fact that correlative conditionals are a type of conditional, and some are the properties of comparatives, but some are peculiar to this construction. Fillmore and McCawley argue that the only way to capture these properties is with a construction-specific rule of some sort (a grammatical construction in the technical sense for Fillmore and a construction-specific transformation for McCawley). McCawley notes that the other alternative is to localize the properties of this construction in some of the words involved. However, as he argues, this is unsatisfactory: An account that localized the properties of the construction in some lexical item (*the*? a comparative operator? some kind of empty category?) would miss the connections that correlative conditionals bear to conditionals, on the one hand, and comparatives on the other. Only in a network of related constructions can one represent both the peculiarities of correlative conditionals and the generalizations that they participate in.

2.4.2 Relative clauses

Sag (1997) develops a hierarchy of phrase types which accounts for a wide variety of relative clauses in English. On his analysis (see also Ginzburg and Sag to appear), the

type *phrase* is elaborated along two dimensions, CLAUSALITY and HEADEDNESS, as shown in Figure 2.31. According to this hierarchy, all phrases are either non-clausal, or belong to one of the four major clause types (imperative, declarative, interrogative or relative).[36] These clause types each dominate a number of subtypes. It is on the supertypes like *inter-cl* that the message value (illocutionary force) is fixed. The *rel-cl* supertype states that all instances express *propositions*, and further that they modify nouns ([MOD [HEAD *noun*]]). The hierarchy in Figure 2.31 also states that every *phrase* is either non-headed (*non-hd-ph*), or belongs to one of the types of headed phrase, such as the *fin-head-subj-phrase* or the *head-filler-phrase*.

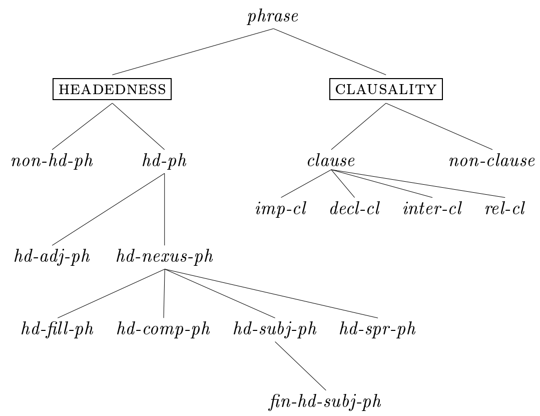


Figure 2.31: Sag's (1997) hierarchy of phrase types

Individual constructions inherit from both sides of this hierarchy, and may or may not add further constraints of their own. In Sag's system, the type *relative-clause* has two immediate subtypes: *wh-relative-clause* and *non-wh-relative-clause*. The distinguishing feature of all *wh* relative clauses is that their non-head daughter contains a *wh* word which is coindexed with the noun the relative clause modifies.

[36] This hierarchy could obviously be extended to include other clause types.

Non-*wh* relative clauses consist of a single slashed head daughter. The item in the slash set of the head daughter is coindexed with the modified noun.

All of the relative clauses considered,[37] inherit from one of the two relative clause types, and some subtype of *headed-phrase*. For example, *wh-subject-relative-clause* (18) inherits from both *wh-relative-clause* and *fin-head-subject-phrase*.

(18) the baker who ate all of the cream puffs

The type *simple-infinitive-relative-clause* (19) inherits from both *non-wh-rel-cl* and *head-comps-phrase*.

(19) the baker to blame for the missing cream puffs

For the details of the analysis, the reader is referred to Sag 1997. What is important here is that with this system Sag is able to capture the generalizations involving all relative clauses and subgeneralizations involving only some types of relative clause, while accounting for such intricate phenomena as pied-piping and the ordering of different types of relative clauses should they modify the same noun. Further, this constructional analysis makes no appeal to phonologically empty relative complementizers assumed in many GB analyses. As such, it is yet another example of a constraint-based analysis obviating the need for a type of empty category.

2.4.3 Summary

To some extent the existence of both constructions and rich lexical entries leaves us with an overabundance of analytical tools. That is, while it is often possible to show that some phenomenon requires the positing of a construction (especially if one is avoiding positing phonologically empty elements), many phenomena could be analyzed as either the effects of an additional (set of) constructions or as some type of lexical alternation. (The exceptions are lexical alternations with morphological expression: These are required to be treated lexically by the Lexical Integrity

[37] Sag does not extend his analysis to free relatives, but notes that such an extension is possible.

Hypothesis of Bresnan and Mchombo 1987, widely accepted by Construction Grammarians.[38]) For example, subject-auxiliary inversion could be a matter of a lexical alternation that places the subject on the complements list (Warner 1993 and others) or a matter of a construction that realizes both the subject and the complements after the auxiliary (Pollard and Sag 1994, Fillmore 1999). Further, constructional analyses, especially those of 'peripheral' phenomena that involve restricted classes of words or other complications, REQUIRE rich lexical entries to represent this information. The situation will be ameliorated somewhat if lexical rules are abandoned in favor of modeling lexical alternations with a type hierarchy. The type-hierarchy system appears to be somewhat more restricted than lexical rules. If so, then a larger class of phenomena may become clearly constructional.

This overabundance of analytical tools is not, however, a defect in the theory. Rather, it is an avenue to new discoveries. That constructions are necessary is shown by the examples discussed above. As discussed in Chapter 6, a construction-based approach to syntax looks to be the most promising for integrating sociolinguistic variation into the theory of grammar. Perhaps it will eventually be possible to develop a set of heuristics for determining when a phenomenon reflects the existence of a construction, and when it reflects a lexical alternation; perhaps both analyses exist in different people's grammars.

2.5 Properties of English auxiliaries

The final aspect of HPSG that I would like to review is the treatment of four major properties of English auxiliaries: negation, inversion, ellipsis and tag questions. Warner (1993, 2000), Kim and Sag (1995), Kim (2000), Bender and Flickinger (1999), and Sag and Wasow (1999) analyze the properties in terms of lexical alternations

[38] In the Construction Grammar literature, morphological rules are also called 'constructions' but crucially obey the Lexical Integrity Hypothesis in that morphological constructions are distinct from syntactic constructions: the constructions that alter the morphological form of a word do not also combine it with other words.

involving the auxiliaries.[39] Inversion has also, however, been analyzed as a special construction or phrase structure schema (Pollard and Sag 1994, Fillmore 1999, Sag 1999). Sag (1999) also proposes a constructional analysis of ellipsis.

In this presentation, I will draw on Warner's (2000) presentation of the negation facts within a multiple inheritance hierarchy and Sag's (1999) alternative constructional account of ellipsis and inversion. I will also update Bender and Flickinger's (1999) analysis of tag questions as a construction related to the inversion and ellipsis constructions. This presentation will focus on the standard variety represented in the literature cited, with some reference to AAVE.

Before turning to the analysis of these properties of auxiliaries it is necessary to first consider what a basic (non-negated, non-inverted, etc.) lexical entry for an auxiliary looks like. As forms of *be* are most relevant here, they will be used to illustrate the whole class.

Following Ross (1967)a and Gazdar et al. (1982), auxiliaries are analyzed as taking VP complements.[40] In the theory assumed here, it follows from this that auxiliaries are raising (or control) verbs. Thus the lexical entry for *be* given in Figure 2.32 is similar in the relevant respects to that given for *tend* in Figure 2.18, page 27.

Be also differs from *tend* in two respects: First, while both *be* and *tend* are verbs, *be* is also an auxiliary verb. They will thus be distinguished by a feature AUX. If all of the auxiliary properties are treated as lexical alternations, it is sufficient to mark *tend* as [AUX −] and *be* as [AUX +]. However, Sag's (1999) constructional analysis reinterprets the feature AUX. Accordingly, I have not given *be* an AUX specification in Figure 2.32 and postpone further discussion of this feature until we get to the analysis of negation below.

More importantly, while *tend* takes an infinitival CP complement, *be* selects for a predicative phrase (coded here as any subject-seeking, complement-saturated constituent that is [PRED +]). Predicative phrases include predicative NPs (20), APs

[39] Lexical alternations have been treated within HPSG variously as lexical rules (Flickinger et al. 1985, Pollard and Sag 1987, Flickinger 1987) and particular interactions of types within a multiple inheritance hierarchy (Riehemann 1999, Kathol 1999, Koenig 1999, Davis and Koenig 2000, Warner 2000).

[40] See also Akmajian et al. (1979) where non-finite auxiliaries take VP complements, although finite auxiliaries still instantiate an Aux node.

$$
\left[\begin{array}{ll}
\multicolumn{2}{l}{\textit{lexeme}} \\
\textsc{phon} & \langle \text{ be } \rangle \\
\textsc{synsem} & \left[\begin{array}{ll}
\textsc{local} & \left[\begin{array}{ll}
\textsc{cat} & \left[\begin{array}{ll}
\textsc{head} & \textit{verb} \\
\textsc{arg-st} & \left\langle \boxed{1}, \left[\textsc{local} \mid \textsc{cat} \left[\begin{array}{ll}
\textsc{head} & \left[\textsc{pred}\ +\right] \\
\textsc{subj} & \langle \boxed{1} \rangle \\
\textsc{spr} & \langle\, \rangle \\
\textsc{comps} & \langle\, \rangle
\end{array}\right]\right]\right\rangle
\end{array}\right] \\
\textsc{cont} & \left[\begin{array}{ll}
\textsc{key} & \boxed{2} \\
\textsc{rels} & \langle \boxed{2} \rangle
\end{array}\right]
\end{array}\right]
\end{array}\right]
\end{array}\right]
$$

Figure 2.32: Lexical entry for *be*

(21), and PPs (22), and progressive and passive participial VPs (23–24).

(20) Kim is a doctor.

(21) Kim is tired.

(22) Kim is at a party.

(23) Kim is sleeping.

(24) Kim was left behind.

Note that predicative NPs differ slightly in their range of possible internal syntactic structure from non-predicative NPs (25).[41] Similarly, not all adjectives can be used predicatively (26).

(25) a. Dana was advisor to the committee.
b. Mary is too much of a fool to take seriously.
c.*I ran after advisor to the committee.
d.*Too much of a fool to take seriously was appointed to the committee.

[41] From Warner 1993:74.

(26)*This suggestion is mere.

Traditionally, this range of complementation patterns has been broken up into the copula (taking NP, AP, and PP complements), progressive *be* and passive *be*. However, Lapointe (1980), Falk (1984), and Warner (1993) all argue for collapsing all of these complementation patterns as in Figure 2.32. Warner gives two main arguments for this position: First, the passive and progressive participles occur without *be*, with progressive and passive 'meaning' (27–28).[42] (In the case of passive, this meaning is actually just an alternative linking of semantic to syntactic arguments.) Therefore, there is reason to attribute passive or progressive 'meaning' to special *be* lexemes that select these participles as complements.

(27) What, Major droning on again!

(28) Anyone thought to be hiding contraband will be interrogated.

Second, all combinations of complement types shown in (20–24) can appear coordinated as the complement of one single *be*. Some of these combinations are illustrated in (29).[43]

(29) a. He is [$_{\text{AP}}$ very angry indeed] and [$_{\text{VP[prog]}}$ throwing furniture about the room].
b. Paul is [$_{\text{PP}}$ already in the car] and [$_{\text{VP[prog]}}$ waiting for you].
c. I'm [$_{\text{VP[prog]}}$ still expecting to go] and [$_{\text{AP}}$ very keen about the prospect].
d. Paul is [$_{\text{AP}}$ horribly misshapen], [$_{\text{NP}}$ a creature of darkness], and [$_{\text{VP[pass]}}$ thought to practice witchcraft]. Please don't ask him round again.
e. The contraband was [$_{\text{PP}}$ inside the wheel arch] and [$_{\text{VP[pass]}}$ thought to be safely hidden].
f. John was [$_{\text{VP[pass]}}$ put into a state of turmoil by the decision] and [$_{\text{VP[prog]}}$ hoping for its reversal].

[42](27) is from Warner 1993:75.
[43]Falk (1984) attributes this type of example to Thomas Wasow. The examples in (29) are from Warner 1993:24.

The same holds true for AAVE, both in sentences with overt forms of *be* and in copulaless sentences. (30) illustrates the claim for copulaless sentences.[44]

(30) a. Paul already in the car and waitin for you.
b. Paul red as a beet and lookin like he wants to/wanna disappear.
c. We hopin to go and too pumped about it.
d. %Paul ugly, a crook, and known to smoke crack. Don't you ever invite him again.

We now turn to the first property of English auxiliaries: sentential negation.

2.5.1 Negation

Sentential negation in English is expressed by *not* immediately following a finite auxiliary, or by special negated forms of the auxiliaries:[45]

(31) a. Kim is not happy.
b. Kim isn't happy.

Sentential negation is to be distinguished from constituent negation (Klima 1964). In (32), the first *not* represents sentential negation, and the second, constituent negation:

(32) Kim [can [not] [not eat donuts]].

Warner (1993), Kim and Sag (1995), Kim (2000) and Warner (2000) argue for an analysis of (uncontracted) sentential negation in which *not* is an argument of the auxiliary. That is, examples like (31) involve auxiliaries like the one sketched in Figure 2.33, which gives rise to constituent structures like the one shown in Figure 2.34.[46]

The evidence for this analysis includes the behavior of *not* in VP ellipsis, the distribution of *do*, and lexical idiosyncrasy in the relative scope of the auxiliary and *not*.

[44](30d) was accepted by two consultants, rated marginal by one, and rejected by one. The consultant who rejected it rejects passives with the null copula in general, and accepts (30d) with overt *is*. The one who rated it as marginal remarked that it is better with a pronominal subject. (30c) was only checked with three consultants, and accepted by two and rated as slightly marginal by the other.
[45]See Zwicky and Pullum 1983 for arguments that *-n't* is an affix and not a clitic.
[46]Again, this is a *lexeme* type. Further information needs to be filled in by inflectional types to make a word that can be used in the syntax.

$$\begin{bmatrix} \textit{lexeme} \\ \text{PHON } \langle \text{ is } \rangle \\ \text{SYNSEM} \mid \text{LOCAL} \begin{bmatrix} \text{CAT} \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \textit{verb} & \\ \text{FORM} & \textit{fin} \end{bmatrix} \\ \text{ARG-ST} & \left\langle \boxed{1}, \text{Adv}[\textit{not}], \begin{bmatrix} \text{LOCAL} \mid \text{CAT} \begin{bmatrix} \text{HEAD} & [\text{PRED} +] \\ \text{SUBJ} & \langle \boxed{1} \rangle \\ \text{SPR} & \langle\, \rangle \\ \text{COMPS} & \langle\, \rangle \end{bmatrix} \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

Figure 2.33: *Be* with sentential negation

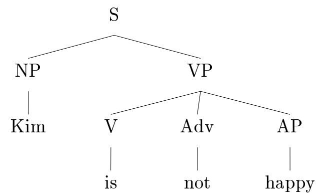


Figure 2.34: Constituent structure of sentential negation

Unlike other VP initial adverbs, *not* may be stranded by ellipsis. Further, *not* may only be stranded by ellipsis when it is preceded by a FINITE auxiliary. That is, only SENTENTIAL negation can be stranded.[47]

(33) a. Kim said he could have heard the news, but Lee said that he could not.
b. *Kim said he could have heard the news, but Lee said that he could have not.
c. *Kim has written a novel, but Lee has never.

In a theory where ellipsis does not involve empty categories, the *not*-as-complement analysis of sentential negation predicts these facts. (33b) and (33c) are out because the modifier (*never* or *not*) has no constituent to attach to. In (33a), on the other hand, *not* is not a (syntactic) modifier, but rather a complement of *could*.

This analysis of sentential negation also provides part of an account of the phenomenon called '*do* support' (Chomsky 1955, 1991). Unstressed (i.e., non-emphatic) *do* occurs in all of the auxiliary constructions, including sentential negation, but not elsewhere:

(34) a. Kim did not leave.
b. Did Kim leave?
c. Sandy didn't consider leaving, but Kim did. (=Kim did consider leaving).
d.*Kim did leave.
e. Contrary to popular belief, Kim DID leave.

This pattern can be broken down into three facts: (1) Main verbs cannot participate in these constructions. (2) When there is no other auxiliary present, *do* is 'inserted'. (3) Unstressed auxiliary *do* does not appear elsewhere. The first two facts are accounted for by restricting sentential negation and the other auxiliary constructions to the class of auxiliary verbs and by positing a semantically vacuous auxiliary *do*. The first move rules out sentences like (35), and the second ensures that the sentences in (34a-c) have the same meanings as (35a-c) would were they grammatical.[48]

[47] Examples from Kim and Sag 1995:310–311.

[48] See Kim 2000 for arguments that this general approach is better than a transformational '*do*-support' approach.

(35) a.*Kim left not.
b.*Left Kim?
c.*Sandy didn't consider leaving, but Kim (did) consider.

Sag (1999) proposes to account for the third property by reinterpreting the feature AUX. AUX is a head feature of verbs. In much previous work, all auxiliaries were specified as [AUX +] and all main verbs being specified as [AUX −]. On Sag's reinterpretation, main verbs are still [AUX −], but most auxiliaries are now [AUX *bool*], that is, underspecified for AUX. The exception is auxiliary *do*, which is specified as [AUX +].[49] Certain phrasal signs make reference to this feature. The inversion, ellipsis, and tag constructions discussed below all require [AUX +] heads. Negation, still treated as a lexical alternation, also only applies to [AUX +] words. However, Sag's *finite-verb-phrase* (a subtype of *head-complement-phrase*) instead constrains the AUX and NEG values to be the same. That is, the head of a finite verb phrase may be [AUX +] just in case it is also negated (see Figure 2.35).[50]

$$\begin{bmatrix} \textit{finite-verb-phrase} & \\ \text{HEAD-DTR} \mid \text{SYNSEM} \mid \text{LOCAL} \mid \text{CAT} \mid \text{HEAD} & \begin{bmatrix} \textit{verb} & \\ \text{AUX} & \boxed{1} \\ \text{NEG} & \boxed{1} \\ \text{INV} & - \end{bmatrix} \end{bmatrix}$$

Figure 2.35: *finite-verb-phrase*, subtype of *head-comps-phrase*

Since most auxiliaries are underspecified for AUX, they can appear in this construction with or without negation. Non-emphatic *do*, on the other hand, is always [AUX +], and as such can only head this construction when it is negated. It is still free to participate in the other constructions (inversion, ellipsis and tag) because they

[49] Auxiliary *do* is not to be confused with main verb *do*, which is always [AUX −]. They also differ in that auxiliary *do* only has finite forms.

(i) Kim didn't do it.
(ii) *Kim should do run.

[50] See the discussion of inversion in §2.5 below for an explanation of the INV feature.

are not subject to the constraint in Figure 2.35. Emphatic *do* is unspecified for AUX, so it can occur as the head of a *finite-verb-phrase* without being negated.

(36) Kim DID leave.

Finally, a lexical analysis of sentential negation is apparently required by the idiosyncratic nature of scopal relations. As first noted by Horn (1972) and explored in detail by Warner (2000), the scope of sentential negation with respect to the semantic relation of modals varies from auxiliary to auxiliary. Assuming the interpretations of the various auxiliaries that Warner does, *can, could,* deontic *may, dare, will, would,* and *need* take wide scope negation, while *must, ought, shall, should* and epistemic *may* and *might* take narrow scope negation. (37) and (38)[51] exhibit wide and narrow scope negation, respectively. Note that the polarity of the tag in (37) forces the sentential negation reading as opposed to the constituent negation (narrow scope) reading. In (38), the ellipsis shows that this is also an instance of sentential negation, even though the negation takes narrow scope.

(37) Paul could not have worked as hard, could he?
[Scope: not(possible)]

(38) — Paul should not have been drinking, should he?
— No, he should not.
[Scope: obligation(not)]

As illustrated in (39) and (40), the auxiliaries exhibit the same scope properties with *-n't*:

(39) Paul couldn't have worked as hard, could he?
[Scope: not(possible)]

(40) Paul shouldn't have been drinking, should he?
[Scope: obligation(not)]

[51] (37–40) are from Warner 2000:176–177.

Warner (2000) also notes that the wide and narrow-neg-scope auxiliaries do not fall into natural semantic classes. He is able to group most of the narrow scope auxiliaries as contributing either a *obligation_rel* or a *may-epistemic_rel*, but even this disjunctive classification has exceptions: *need*, which is covered by this classification belongs in fact to the wide-scope class.[52] Treating the relative scope of *not* and the auxiliaries as a lexical matter makes such lexical idiosyncrasies expected. Further, it allows for the constraints concerning scope to be stated fairly simply, as shown below.

Before turning to the remaining auxiliary properties, I will briefly present a formalized version of Warner's analysis. This version departs somewhat in its details from Warner's own formalization, but remains the same in spirit.

Because the scopal relations for an auxiliary remain the same under both forms of negation, Warner proposes a cross-classification, shown in Figure 2.36.[53,54]

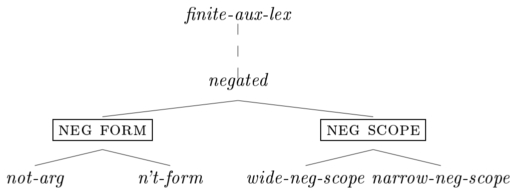


Figure 2.36: Negation subhierarchy

The type *negated* specifies semantic negation. The types under $\boxed{\text{NEG SCOPE}}$ specify the relative scope of negation and the auxiliary, while the types under $\boxed{\text{NEG FORM}}$ specify whether the negation is realized as a *not* complement or as the affix *n't*.

Where I begin to depart from Warner's formalization is in the location of this subhierarchy within the larger hierarchy. For Warner, these are all *lexeme* types. However, the morphology of the *-n't* forms (especially *isn't, wasn't, hasn't*, etc.) show

[52] Warner (arbitrarily) includes *be*, *have*, and *do* in the wide-neg-scope class.

[53] For Warner, this cross-classification is embedded in a larger one, in which all auxiliaries are classified along two dimensions, $\boxed{\text{NEGATION}}$ and $\boxed{\text{INVERSION}}$. Since inversion is treated constructionally here, the hierarchy shown in Figure 2.36 is actually just a subhierarchy of Warner's.

[54] The dashed line connecting *finite-aux-lex* and *negated* indicates that there are intermediate subtypes.

that the type that adds *-n't* can't be a lexeme type: the stem that *-n't* attaches to is a fully inflected word. Warner doesn't go into inflection, but placing the *n't-form* type within the *lexeme* subhierarchy presupposes a simplified view of (inflectional) morphology that almost works for English. On this view, there is a feature STEM in addition to PHON, and lexemes specify only STEM values. Then inflectional types, such as *3sg-present-verb*, specify a relationship between the STEM and PHON values. As Koenig (1999) argues at length, for morphological systems more complex than English inflectional morphology, a constituent structure approach is necessary.[55] In Koenig's model, morphological constituent structure is modeled by types that have a DAUGHTER feature which takes other lexemes as its value.[56]

The most general part of Koenig's hierarchy of signs is given in Figure 2.37.[57] As usual, there are two main subtypes of *sign*: *phrase* and *lexeme*. This latter type is usually named *word*, but Koenig reserves the name *word* for a subtype of *lexeme*, namely those signs that are syntactic atoms—fully inflected words that can go 'out' into the syntax. The other subtypes of *lexeme* are *root* and *stem*. Signs of type *stem* are morphologically complex, but not syntactic atoms. That is, they are incomplete words. Both *words* and *stems* are subtypes of *complex-lexeme*, so they both inherit the specification in Figure 2.38. Since the value of the DAUGHTER feature is of type *lexeme*, the embedded sign can be a *word, stem* or *root*. Signs of type *root* are morphological atoms, that is, they are monomorphemic. This type corresponds to the (implicit) type *lexeme* in Warner's (2000) system. That is, Warner's *negated* subhierarchy would be within the *root* subhierarchy. However, the *root* subhierarchy includes much cross-classification (including the part of speech and valence dimensions), but no morphologically complex types.

[55] Koenig also argues that a constituent structure approach alone isn't sufficient. His model incorporates both multiple inheritance hierarchies and morphological constituent structure.

[56] More precisely, Koenig's *complex-lexemes* have a μ-STRUC feature which takes a value of type μ-*struc*, and μ-*strucs* have a DAUGHTER feature. Since the rest of the μ-STRUC doesn't concern us here, I will use the simpler feature geometry in Figure 2.38.

[57] The $\boxed{\text{MORPH-PHON}}$ partition contains types that model phonological alternations. This aspect of Koenig's theory is not relevant to the discussion at hand.

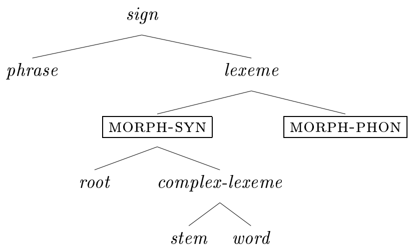


Figure 2.37: Koenig's hierarchy of signs

$$\begin{bmatrix} \textit{complex-lexeme} & \\ \textsc{daughter} & \textit{lexeme} \end{bmatrix}$$

Figure 2.38: Constituent structure of complex lexemes

As noted above, the affix *n't* attaches to fully inflected words. As such, it seems best to model it as a subtype of *word*.[58] This is shown in Figure 2.39. The type *negated* subsumes two partitions or dimensions. This means that any word belonging to the type *negated* must belong to one type from each of $\boxed{\text{NEG FORM}}$ and $\boxed{\text{NEG SCOPE}}$.

Figure 2.40 shows the constraints on the type *negated*.[59] The first thing to note is that this type is only compatible with daughters that are finite non-negated auxiliaries. The mother will also be a finite auxiliary, but specified as [NEG +]. The differing specifications for NEG on the mother and daughter prevent the daughter from also being of type *negated*, i.e., block this type from 'applying' recursively to its

[58]It is tempting to preserve the simpler STEM/PHON model of morphology for English by listing out all of the inflected forms of auxiliaries as separate lexemes. After all, *be* is irregular and as such needs all of its forms listed somehow anyway. However, the are many different uses of *be*, several of which may need their own lexeme types. It would therefore be best if the various forms of *be* were modeled with inflectional types that could 'apply' to all of the different lexemes *be*. In this case, the STEM/PHON system won't work: The inflectional types would already fix the relationship between STEM and PHON, and the *n't-form* type couldn't affect the PHON value further.

[59]For all of these *word* subtypes, I am assuming that a good deal of information, including much of the CONTENT values, is 'copied up' from the daughter to the mother. Ideally, such identities could be stated on a supertype, perhaps as high as *word* itself.

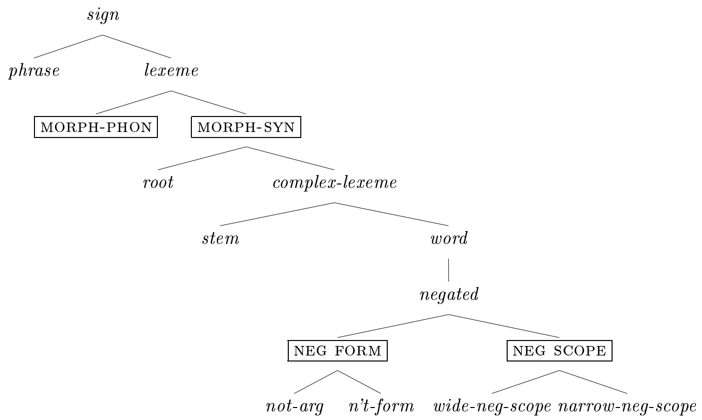


Figure 2.39: Integrating auxiliary negation into Koenig's hierarchy of signs

own output. Note that this means that the type can't be headed, as the head values conflict.[60]

The type *negated* also provides the adverb *not*. However, *not* is not placed directly on the argument structure list. Taking advantage of Koenig's (1999) rearrangement of the ARG-ST feature, the *not* is placed on the ADD-ARG list. The intuition behind these features is that SEM-ARG contains the semantically selected arguments, ADD-ARG contains any additional arguments (such as the expletive subject in the extraposition type), and ARG-LIST is some combination of the two. ARG-LIST is also what constraints such as the argument realization principle should refer to.[61] In Figure 2.40, the SEM-ARG value of the mother is identified with the ARG-LIST value of the daughter. It wouldn't do to put the *not* on the ARG-LIST at this point, because *n't-form* in fact requires that the *not* not appear on the ARG-LIST. However, it is advantageous to introduce *not* on the supertype *negated*, as that way the scope types

[60]It may be possible to make the type headed anyway, with appropriate use of defaults.

[61]Outside of this subsection, I will return to the more familiar ARG-ST notation, with the understanding that this is just an abbreviation for ARG-ST | ARG-LIST.

$$\begin{bmatrix} \textit{negated} & \\ \text{SYNSEM} & \begin{bmatrix}\text{LOCAL} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix}\text{FORM} & \textit{fin}\\ \text{AUX} & + \\ \text{NEG} & +\end{bmatrix} \\ \text{ARG-ST} & \begin{bmatrix}\text{SEM-ARG} & \boxed{A}\\ \text{ADD-ARG} & \langle \text{Adv}[\textit{not}] \rangle\end{bmatrix}\end{bmatrix} \\ \text{CONT} & \begin{bmatrix}\text{KEY} & \boxed{1}\end{bmatrix}\end{bmatrix}\end{bmatrix} \\ \text{DAUGHTER} & \begin{bmatrix} \textit{word} & \\ \text{SYNSEM} \mid \text{LOCAL} & \begin{bmatrix}\text{CAT} & \begin{bmatrix}\text{HEAD} & \begin{bmatrix}\text{FORM} & \textit{fin}\\ \text{AUX} & +\\ \text{NEG} & -\end{bmatrix}\\ \text{ARG-ST} & \begin{bmatrix}\text{ARG-LIST} & \boxed{A}\end{bmatrix}\end{bmatrix}\\ \text{CONT} & \begin{bmatrix}\text{KEY} & \boxed{1}\end{bmatrix}\end{bmatrix}\end{bmatrix}\end{bmatrix}$$

Figure 2.40: Constraints on the type *negated*

can also refer to this element.

Figures 2.41 and 2.42 give the constraints on the types *n't-form* and *not-arg* respectively. The type *n't-form* identifies the SEM-ARG and ARG-LIST values, effectively keeping *not* off the ARG-LIST. Since the word *not* won't appear in the phrase structure anywhere, its semantic contribution has to be incorporated some other way. *N't-form* accomplishes this by adding the RELS value of *not* ($\boxed{C}$) to its own RELS value.[62] The phonological changes are modeled here as a function $f_{n't}$, which takes the phonology of the daughter as its input. The simpler PHON value $\boxed{D} \oplus \langle$ n't $\rangle$ would give the wrong result for *don't, won't, mustn't* and *shan't*. Further, given the function $f_{n't}$, one can stipulate that it is undefined for *am* (cf. **amn't*).

$$\begin{bmatrix} \textit{n't-form} & \\ \text{PHON} & f_{n't}(\boxed{D}) \\ \text{SYNSEM} & \begin{bmatrix}\text{LOCAL} & \begin{bmatrix}\text{CAT} & \begin{bmatrix}\text{ARG-ST} & \begin{bmatrix}\text{SEM-ARG} & \boxed{A}\\ \text{ADD-ARG} & \langle \begin{bmatrix}\ldots \mid \text{RELS} & \boxed{C}\end{bmatrix} \rangle \\ \text{ARG-LIST} & \boxed{A}\end{bmatrix}\end{bmatrix}\\ \text{CONT} & \begin{bmatrix}\text{RELS} & \boxed{B} \oplus \boxed{C}\end{bmatrix}\end{bmatrix}\end{bmatrix}\\ \text{DAUGHTER} & \begin{bmatrix}\text{PHON} & \boxed{D}\\ \text{SYNSEM} & \begin{bmatrix}\text{LOCAL} \mid \text{CONT} \mid \text{RELS} & \boxed{B}\end{bmatrix}\end{bmatrix}\end{bmatrix}$$

Figure 2.41: Constraints on the type *n't-form*

The type *not-arg* (Figure 2.42) is somewhat simpler. Since it has no phonological effects, the PHON values of the mother and daughter are identified ($\boxed{B}$). Further, as *not* will appear as an independent word in the phrase structure when auxiliaries inheriting from this type are used, nothing special need be said about the semantics of *not*. The type *not-arg* does, however, place the *not* onto the ARG-LIST of the mother ($\boxed{3}$).

[62] $\boxed{C}$ in Figure 2.41 must be the RELS value of *not* because *n't-form* inherits from *negated* the constraint [ADD-ARG $\langle$ Adv[*not*] $\rangle$].

$$\begin{bmatrix} \textit{not-arg} & \\ \text{PHON} & \boxed{B} \\ \text{SYNSEM} & \begin{bmatrix}\text{LOCAL} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{ARG-ST} & \begin{bmatrix} \text{SEM-ARG} & \langle \boxed{2} \rangle \oplus \boxed{A} \\ \text{ADD-ARG} & \langle \boxed{3} \rangle \\ \text{ARG-LIST} & \langle \boxed{2}, \boxed{3} \rangle \oplus \boxed{A} \end{bmatrix}\end{bmatrix} \\ \text{CONT} & \begin{bmatrix}\text{RELS} & \boxed{C}\end{bmatrix}\end{bmatrix}\end{bmatrix} \\ \text{DAUGHTER} & \begin{bmatrix}\text{PHON} & \boxed{B} \\ \text{SYNSEM} \mid \text{LOCAL} \mid \text{CONT} \mid \text{RELS} & \boxed{C}\end{bmatrix}\end{bmatrix}$$

Figure 2.42: Constraints on the type *not-arg*

Figures 2.43 and 2.44 show the constraints on *wide-neg-scope* and *narrow-neg-scope* respectively. Both pick up the relevant features the possibility that something scopes in between the handle of the *not_rel* from the ADD-ARG value,[63] as this is the only place it is uniformly available in both *n't-form* and *not-arg*. The different scope relationships are represented as one relation taking the other as its argument. In *wide-neg-scope*, the handle of the auxiliary's main relation ($\boxed{1}$) is identified with the argument of the *not_rel*. In *narrow-neg-scope*, the handle of the *not_rel* is identified with the argument of the auxiliary. Further, the argument of the *not_rel* is identified with the handle of the complement of the auxiliary ($\boxed{2}$).[64] The KEY specifications are meant to constrain each type to the relevant auxiliaries. As noted by Warner, these classes still have exceptions, which will need to be listed in some way.

[63] The '...' before KEY in these avms indicate an abbreviated feature path.

[64] This is still abstracting away from the possibility of quantifiers. In fact, the relevant HANDLE values are constrained to be out-scoped by or equal to the relevant ARG values, to allow for the possibility that some quantifier scopes in between. (This is why handles, and not indices, must be used for this purpose. For details, see Copestake et al. 1999.) This also means that the constraints on *narrow-neg-scope* do not contradict the constraint (on the base entries for the auxiliaries) that the handle of the auxiliary's complement is out-scoped by or equal to the argument of the auxiliary's relation.

$$\begin{bmatrix} \textit{wide-neg-scope} & \\ \text{SYNSEM} \mid \text{LOCAL} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{ARG-ST} \mid \text{ADD-ARG} & \left\langle \begin{bmatrix} \ldots \mid \text{KEY} & \begin{bmatrix}\text{ARG} & \boxed{1}\end{bmatrix}\end{bmatrix} \right\rangle \end{bmatrix} \\ \text{CONTENT} & \begin{bmatrix} \text{HANDLE} & \boxed{1} \\ \text{KEY} & \begin{bmatrix} \neg\ (\textit{oblig\_rel} \vee \textit{may-epistemic\_rel}) \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

Figure 2.43: Constraints on the type *wide-neg-scope*

$$\begin{bmatrix} \textit{narrow-neg-scope} & \\ \text{SYNSEM} \mid \text{LOCAL} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{ARG-ST} \mid \text{ADD-ARG} & \left\langle \begin{bmatrix} \ldots \mid \text{KEY} & \begin{bmatrix}\text{HANDLE} & \boxed{1} \\ \text{ARG} & \boxed{2}\end{bmatrix}\end{bmatrix} \right\rangle \\ \text{SEM-ARG} & \left\langle [\ ], \begin{bmatrix} \ldots \mid \text{KEY} & \begin{bmatrix}\text{HANDLE} & \boxed{2}\end{bmatrix}\end{bmatrix} \right\rangle \end{bmatrix} \\ \text{CONTENT} & \begin{bmatrix} \text{KEY} & \begin{bmatrix} \textit{oblig\_rel} \vee \textit{may-epistemic\_rel} \\ \text{ARG}\ \boxed{1} \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

Figure 2.44: Constraints on the type *narrow-neg-scope*

Ain't

At this point, a few words are in order about the status of *ain't* in AAVE and how it fits in with the analysis of sentential negation reviewed here. Weldon (1994) argues, on the basis of patterns of variation as well as specific attested utterances, that *ain't* serves as an inflected form of the auxiliaries *be*, *have*, and past tense *do*. Examples (41–43), from Weldon 1994:362–363 illustrate each of these use of *ain't*.

(41) They ain't 'posed to know we taping.

(42) He ain't started my grill yet.

(43) I ain't give you none, Boo, did I?

The alternative analysis Weldon argues against is Debose's (1994) proposal that *ain't* (as well as *not*) is a tense/aspect neutral auxiliary, separate from *be*, *have*, and *do*.

The syntactic evidence that Weldon adduces for her claim is as follows. First, *ain't* can appear in matrix polar questions and tag questions—environments that require a tensed auxiliary.[65] In this, it contrasts with *not*, although *ain't* and *not* might appear to occupy the same position in (46).

(44) a. Ain't that little boy cute?
b. *Not that little boy cute?

(45) a. That stuff is still in the refrigerator, ain't it?
b. *That stuff is still in the refrigerator, not it?

(46) a. They ain't best friends.
b. They not best friends.

Secondly, when *ain't* is used as a past tense form of *do*, it gets the matching tag, as shown in (43) above. Tags that might go with other uses of *ain't* won't work here (Weldon 1994:387).

(47)*I ain't give you none, Boo, do I?/am I?/have I?

[65] (44) and (45) are from Weldon 1994:378.

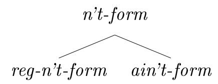


Figure 2.45: Two subtypes of *n't-form*

It's true that forms of *do* are the default for tags in the sense that tags on sentences headed by main verbs always take *do*. However, if *ain't* were truly a tense/aspect neutral auxiliary, one would expect it to be able to take *have*, *be* or present tense *do* tags regardless of the 'use' it was being put to in the matrix clause.

Finally, if *ain't* is indeed a form of *do* only in the past tense,[66] then it is not, in fact, completely neutral with respect to tense.

The most straightforward way to include these facts in the analysis of negation reviewed here is through the function $f_{n't}$. This is the function that determines the phonological form of auxiliaries with contracted *n't*. Even without *ain't*, this function must accommodate idiosyncratic phonology, e.g., *won't*/**willn't* (Zwicky and Pullum 1983). It can easily be extended to allow *ain't* as a possible output form for the inputs *is, are, has, have* and *did*.[67]

Another possibility is that the type *n't-form* in Figure 2.41 has two subtypes, as shown in Figure 2.45. On this conception, the type *n't-form* would leave its PHON value unspecified. The type *reg-n't-form*[68] would be specified as [PHON $f_{n't}(\boxed{\text{D}})$], assuming the same $f_{n't}$ as the earlier *n't-form* did. The type *ain't-form* would instead be specified as [PHON *ain't*]. This type would have to require feature specifications on its DAUGHTER that restrict it to *be*, *have* and past tense *do*. Alternatively, the PHON value of *ain't-form* could be given by a function $f_{ain't}$. $f_{ain't}$(have), etc., would give *ain't*, while $f_{ain't}$(would), etc., would be undefined. Note that, since this is a morphological type, positing one type that can produce the *ain't* forms of *have, be* and *do* is not the same as positing a tense/aspect neutral auxiliary *ain't*. The *ain't*

[66] Weldon (1994) finds no examples of *ain't* in present-tense *do* contexts in her data.
[67] Such morphological functions are not presumed to provide a unique output for every input, but rather can give two or more variant outputs for one input.
[68] So named because the relationship between *will* and *won't* is at least more predictable than the relationship between *has* and *ain't*.

form of *have* would retain the semantic and morphosyntactic features of *have* (with the exception of the person/number features), and likewise for the *ain't* forms of *be* and *do*.

The advantage of positing the type *ain't-form* over dealing with *ain't* within the function $f_{n't}$ is that *ain't* is not merely an alternative form of *hasn't/isn't*, etc. It is highly marked in its social/stylistic value. Consider, for example, this description of one use of *ain't* from Anna Quindlen's column entitled "It's the Cult of Personality" in *Newsweek*, August 14, 2000.

> Brokaw and Bush, two guys just standing around talking. Shirtsleeves. Sunshine. Fence posts. You get the idea. The candidate was hunkered down at the ranch, going *mano a mano* for a couple of endless, empty minutes with the anchorman. The governor showed off his Yiddish—"kibitzing" was how he described what he was doing with his father, the former president—but the seminal moment was in fluent good-ole-boy. "I know you are a pretty good fisherman," Bush said to Brokaw, who was angling for the name of the as yet unknown vice-presidential nominee. "Yes, you are, but I ain't catching."
>
> Take a good look at that verb, fellow voters, and consider what the meaning of "ain't" ain't. It ain't good English, of course, and it ain't necessarily an entirely natural locution for a graduate of Andover and Yale, even by way of west Texas. What it is is a marker for the most important issue of this election. Relaxed, a little irreverent, down-home: that "ain't" is supposed to communicate a whole tractor load of material about the Bush personality.

Of course, this is a description of the social and stylistic value of *ain't* in a certain kind of white speech. There's no reason to expect it to be the same in AAVE. However, given the stigmatization and stereotyping of *ain't* it would be surprising for it to have no social or stylistic value in AAVE. If, as argued in Chapter 6, social value is indeed a matter of grammar, the type *ain't-form* provides a place to record it.

2.5.2 Inversion

The next property of English auxiliaries is inversion. Following Pollard and Sag (1994), Fillmore (1999) and Sag (1999), I adopt here a constructional analysis of

subject-auxiliary inversion. On this analysis, inverted structures are built by a separate phrase structure rule that realizes both the subject and the complement(s) after the auxiliary. A preliminary version of the specific rule assumed here (to be revised in the discussion of ellipsis below) is given in Figure 2.46.[69]

$$
\begin{bmatrix}
\textit{sai-phrase} & \\
\text{PHON} & \boxed{B} \oplus \textit{list} \\
\text{SYNSEM} & \begin{bmatrix}\text{LOC} \mid \text{CAT} & \begin{bmatrix}\text{SUBJ} & \langle\,\rangle \\ \text{COMPS} & \langle\,\rangle\end{bmatrix}\end{bmatrix} \\
\text{HD-DTR} & \begin{bmatrix}\text{PHON} & \boxed{B} \\ \text{SYNSEM} & \begin{bmatrix}\text{LOC} \mid \text{CAT} & \begin{bmatrix}\text{HEAD} & \begin{bmatrix}\textit{verb} & \\ \text{INV} & + \\ \text{FORM} & \textit{fin}\end{bmatrix} \\ \text{SUBJ} & \langle \boxed{1} \rangle \\ \text{COMPS} & \textbf{sts}(\boxed{A})\end{bmatrix}\end{bmatrix}\end{bmatrix} \\
\text{NON-HD-DTRS} & \langle \boxed{1} \rangle \oplus \boxed{A}
\end{bmatrix}
$$

Figure 2.46: Subject-auxiliary inversion construction (preliminary version)

The crux of this analysis is the feature INV, introduced by Gazdar et al. (1982) and further justified in Gazdar et al. (1985). The *sai-phrase* requires an [INV +] element as its head. All main verbs are specified as [INV −]. It follows that the *sai-phrase* will always be an auxiliary and the construction need not specify [AUX +]. Both INV and AUX are necessary features, however: While most auxiliaries are underspecified for the feature INV, there are exceptions both ways. The auxiliary *better* (48) is [INV −], while 1st singular *aren't* (49) is [INV +] (Sag 1999).[70]

(48) a. You better not do that.
b.*Better I do that?.

[69] This construction is a subtype of *headed-phrase* and, as such, inherits the constraints shown in Figure 2.13, page 21. The constraint on the PHON value of this construction interacts with the constraints it inherits from *headed-phrase* to ensure that the PHON value of an *sai-phrase* is the PHON value of the head daughter followed by the PHON values of the non-head daughters (i.e., the subject and then the complements).

[70] Sag also cites futurate *shall* and *mightn't* with wide scope negation as [INV +] forms.

(49) a. Aren't I lucky?
b.*I aren't lucky.

In order to capture the facts in (49), and to block VPs (as opposed to lexical verbs) from heading the *sai-phrase* (50), the *finite-verb-phrase* construction must constrain its head to be [INV −] (cf. Figure 2.35, page 53).

(50)*[[Is happy] Kim]?

Note that subject-auxiliary inversion is not limited to matrix polar questions. In a more complete grammar, this *sai-phrase* would be a supertype to a polar question construction, an inverted conditional construction (51), and a blessings/curses construction (52) (Fillmore 1999).[71]

(51) Had we known better, we would never have come.

(52) a. May all your wildest dreams come true.
b. May your teeth fall out on your wedding day.

Likewise, it is involved in the negative inversion constructions in both Standard English (Fillmore 1999) and AAVE (Labov et al. 1968, Sells et al. 1996).[72]

(53) Never have we seen such a sight.

(54) Can't nobody say nothin to dem peoples.

[71] These constructions, of course, impose further constraints, including the restriction to certain auxiliaries for the inverted conditional construction (i) and the blessings/curses construction (ii).

(i) *Did we know better, we would never have come.
(ii) Will all your teeth fall out.

While grammatical, (ii) is not an instance of the blessings/curses construction.

[72] (54) is a naturally occurring example collected by Sells et al. (1996):592.

2.5.3 Ellipsis

Following Sag (1999), I adopt a constructional analysis of ellipsis. The description of *ellipsis-phrase* is given in Figure 2.47. This construction, a subtype of *headed-phrase*, is similar to the *head-complements-phrase* in that it (potentially) realizes some complements, but passes up the subject requirement. However, the relationship between the COMPS specification of the head daughter and the NON-HD-DTRS value is less direct than in the *head-complements-phrase*. Here, the NON-HD-DTRS ($\boxed{B}$) correspond to whatever is left of the COMPS list after some (non-zero) number of complements are elided (*nelist*(*elided*)).

$$\begin{bmatrix} \textit{ellipsis-phrase} & \\ \text{SS} \mid \text{LOC} \mid \text{CAT} & \begin{bmatrix}\text{SUBJ} & \boxed{A}\\ \text{COMPS} & \langle\,\rangle\end{bmatrix} \\ \text{HD-DTR} \mid \text{SS}\mid\text{LOC}\mid\text{CAT} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix}\textit{verbal} & \\ \text{AUX} & +\end{bmatrix}\\ \text{SUBJ} & \boxed{A}\\ \text{COMPS} & \textbf{sts}(\boxed{B}) \oplus \textit{nelist}(\textit{elided})\end{bmatrix}\\ \text{NON-HD-DTRS} & \boxed{B}\end{bmatrix}$$

Figure 2.47: Ellipsis construction

That ellipsis can affect some but not all of the complements of an auxiliary is shown by the examples in (55–58), from Warner 2000:205.[73]

(55) Are there any first-year students angry about their grades?
—Yes, there are some.
—No, there aren't any.

It is also possible to leave out the first complement of the *there*-copula, as shown in (56–58), from Warner 2000:205–206.

(56) Are there any first-year students angry about their grades?
—No, but there are upset with their teachers.

[73] These examples all involve non-locative second complements. As Warner shows, the locative second complements appear to be optional, so that eliding them doesn't place the same requirements on the preceding linguistic context that ellipsis usually does.

(57) He didn't tell me there was any treasure buried in the garden, but I wonder if there is hidden in the orchard.

(58) There is a complex variable hidden in the first formula, so I wonder if there is lurking in the second.

Warner takes these to be examples of ellipsis, but it is not clear how one could distinguish examples like (56) through (58) from examples of pseudo-gapping, like (59), from Miller (1992:94).

(59) John spoke to Mary more often than Peter did to Ann.

Accordingly, I will assume that (56–58) are instances of a separate construction (pseudo-gapping), and that ellipsis can only affect a right-most sublist of the COMPS list.

However, as shown in (60), *not* cannot be elided, even when it appears as the complement of an auxiliary and everything after it is elided.

(60) a. Kim likes bagels, but Sandy does not.
b. Kim does not like bagels, and Sandy does not, either.
c. #Kim likes bagels, but Sandy does.
d. #Kim does not like bagels, and Sandy does, either.

To capture this fact, the elided elements in an instance of *ellipsis-phrase* are unified with the type *elided*, shown in Figure 2.48. Since *not* is arguably the only word selected as a complement by English auxiliaries, the ellipsis of *not* can be blocked by ensuring that *elided* only unifies with the *synsem*s of phrases. The constraints shown in Figure 2.48 accomplish this by taking advantage of the fact that the CAT values of words and phrases are already distinguished. Since ARG-ST is only appropriate for words,[74] we need the subhierarchy in Figure 2.49. Since *phrasal-cat* and *lexical-cat* are sister types in this hierarchy, they are incompatible. *Not*, as a word, will have a CAT value of type *lexical-cat* and its *synsem* won't be able to unify with the type *elided*.

[74] This is crucial to the theory of locality of selection: ARG-ST values, unlike SUBJ/SPR/COMPS values don't get canceled off. If they were to project up the tree, words could select for properties such as the case of the complement of a complement.

$$\begin{bmatrix} \textit{elided} & \\ \text{LOCAL} & \begin{bmatrix} \text{CAT} & \textit{phrasal-cat} \\ \text{CONT} & \ldots \end{bmatrix} \end{bmatrix}$$

Figure 2.48: Constraints on the type *elided*

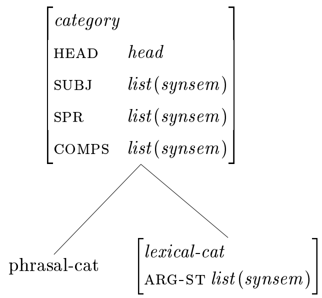


Figure 2.49: Subhierarchy of *category* types

There will also be semantic constraints, on the type *elided* (' ... ' in Figure 2.48) and/or on the type *ellipsis-phrase*, that model the interpretation of the elided phrases and their dependence on preceding linguistic context. Finally, the type *elided* is taken to be a subtype of *non-canonical* synsem, similar to the type *gap* (see Figure 2.50).

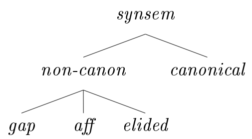


Figure 2.50: Subhierarchy of *synsem* types

The type *ellipsis-phrase* only produces ellipsis in non-inverted clauses. However, ellipsis is also possible in inverted clauses, as shown in (61).

(61) Everyone else has pledged their support. Will you?

As with non-inverted clauses, ellipsis affects a right-most sublist of the complements of the auxiliary:

(62) There aren't many second-year students angry about their grades. Are there any first-years?

It turns out that these can be modeled by making a change to the existing *sai-phrase*. As shown in Figure 2.51, all that is required is that the COMPS value of the head daughter be similar to that of the head daughter in the *ellipsis-phrase*. They differ, however, in that the list of elided synsems in the *sai-phrase* may be the empty list. This allows this single *sai-phrase* to model both elliptical and non-elliptical inverted clauses.

The ability to model both elliptical and non-elliptical inverted clauses with one type is especially desirable given that ellipsis is possible in more than one subtype of *sai-phrase*.[75] (For more on the subtypes of *sai-phrase*, see Fillmore 1999.)

[75] If the semantic constraints on ellipsis involve the whole phrase and not just the type *elided*, this *sai-phrase* may in fact need to be separated into two, one for ellipsis and one for other inversion constructions. However, with lexical amalgamation of relations, it should be possible to state all of the necessary conditions on the type *elided*.

$$\begin{bmatrix} \textit{sai-phrase} & \\ \text{SS} \mid \text{LOC} \mid \text{CAT} & \begin{bmatrix} \text{SUBJ} & \langle\ \rangle \\ \text{COMPS} & \langle\ \rangle \end{bmatrix} \\ \text{HD-DTR} \mid \text{SS} \mid \text{LOC} \mid \text{CAT} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \textit{verb} & \\ \text{INV} & + \\ \text{FORM} & \textit{fin} \end{bmatrix} \\ \text{SUBJ} & \langle \boxed{2} \rangle \\ \text{COMPS} & \textbf{sts}(\boxed{A}) \oplus \textit{list}(\textit{elided}) \end{bmatrix} \\ \text{NON-HD-DTRS} & \langle \boxed{2} \rangle \oplus \boxed{A} \end{bmatrix}$$

Figure 2.51: Subject-auxiliary inversion construction (final version)

(63) a. Had we known what was in it, we wouldn't have opened it.
b. —There was no way to know what was in it.
—Yes, but had we, we wouldn't have opened it.

2.5.4 Tag questions

Bender and Flickinger (1999) analyze tag questions by means of a lexical rule that produces the auxiliaries that head tags. The constructional analysis of ellipsis and inversion developed here allows for an analysis of tags as produced by a construction, and in particular, a subtype of *sai-phrase*.

It may appear that tag questions such as (64) are just special uses of elliptical inverted sentences.

(64) Bob left early today, didn't he?

However, there are at least two kinds of evidence to support a separate tag construction. First, the subject of a tag must be a pronoun. Another referentially dependent expression, such as an epithet, won't do:

(65)*Bob left early, didn't the slacker?

Second, the subject of the tag is always coindexed with the subject of the clause it modifies, and not any other, potentially more salient, NP:

(66) a. There were three people there, weren't there?
b.*There were three people there, weren't they?

Figure 2.52 shows the constraints on the type *tag-phrase*. As a subtype of *sai-phrase*, it inherits all of the constraints in Figure 2.51.

$$\left[\begin{array}{ll}
\multicolumn{2}{l}{\textit{tag-phrase}} \\
\text{SS} \mid \text{LOC} & \left[\begin{array}{ll}
\text{CAT} & \left[\begin{array}{ll}
\text{HEAD} \mid \text{MOD} & \left[\begin{array}{ll}
\text{CAT} & \left[\begin{array}{ll}
\text{HEAD} & \left[\begin{array}{ll} \multicolumn{2}{l}{\textit{verb}} \\ \text{AGR} & \boxed{5} \end{array}\right] \\
\text{SUBJ} & \langle\, \rangle
\end{array}\right] \\
\text{CONT} & \left[\begin{array}{ll}
\text{MSG} & \left[\begin{array}{ll} \multicolumn{2}{l}{\textit{propositional}} \\ \text{SOA} & \boxed{6} \end{array}\right] \\
\text{KEY} & \boxed{2} \\
\text{TENSE} & \boxed{3}
\end{array}\right]
\end{array}\right]
\end{array}\right] \\
\text{CONT} & \left[\begin{array}{ll}
\text{MSG} & \left[\begin{array}{ll} \multicolumn{2}{l}{\textit{tag_rel}} \\ \text{SOA} & \boxed{6} \end{array}\right]
\end{array}\right]
\end{array}\right] \\
\text{HD-DTR} \mid \text{SS} \mid \text{LOC} & \left[\begin{array}{ll}
\text{CAT} & \left[\begin{array}{ll}
\text{HEAD} & \left[\text{AGR} \ \boxed{1}\right] \\
\text{SUBJ} & \left\langle \boxed{4} \left[\begin{array}{ll} \multicolumn{2}{l}{\textit{pron-synsem}} \\ \text{LOC} & \left[\begin{array}{ll} \text{CAT} \mid \text{HEAD} \mid \text{AGR} & \boxed{1} \\ \text{CONT} \mid \text{INDEX} & \boxed{5} \end{array}\right] \end{array}\right] \right\rangle
\end{array}\right] \\
\text{CONT} & \left[\begin{array}{ll} \text{KEY} & \boxed{2} \\ \text{TENSE} & \boxed{3} \end{array}\right]
\end{array}\right] \\
\text{NON-HD-DTRS} & \langle \boxed{4} \rangle \oplus \textit{list}(\text{Adv}[\textit{not}])
\end{array}\right]$$

Figure 2.52: Constraints on the type *tag-phrase*

The first thing to note is that, in addition to the constraints relating the COMPS value of the head daughter to the NON-HD-DTRS value inherited from *sai-phrase*, the *tag-phrase* further stipulates that only complements that are the word *not* may be realized. This allows the somewhat stilted sentences in (67) while correctly ruling out (68).

(67) a. Bob left, did he not?
b. There are several people angry about this, are there not?

(68) a.*Bob left, didn't he leave?
b.*There are several people angry about this, are there not several people?

Thus this construction solves a problem left unresolved by the analysis of Bender and Flickinger (1999): namely, the possibility of retaining complement *not* in tags. The other open question, how to model the polarity effects shown in (69) still awaits a theory of polarity phenomena for its resolution.

(69) a. Sara slept, didn't she?
b. Sara didn't sleep, did she?
c. Sara slept, did she? (challenge tag)
d.*Sara didn't sleep, didn't she?

The rest of the constraints on this type are analogous to those posited in Bender and Flickinger 1999 and motivated there. They will be only briefly reviewed here.

The MOD specification indicates that tag phrases modify declarative sentences. This accounts for their distribution, and also serves as a means to constrain the tag to agree with the sentence it modifies in several ways.[76] In particular, the auxiliary in the tag must match the auxiliary in the sentence, or, if the sentence is headed by a main verb, be a form of *do*. On Bender and Flickinger's analysis, this is done by matching the KEY values of the tag auxiliary and the modified sentence.[77] Tags also agree with their sentences in tense, modeled here by a feature TENSE.[78]

[76]The feature MSG (message) is a representation of illocutionary force. A verb can select for a declarative clause by constraining its complement to be [MSG *propositional*]. The semantic contribution of the tag is also represented as a message, or type of illocutionary force. For more on clause types and their semantics, see Ginzburg and Sag to appear.

[77]To make this work, Bender and Flickinger assume that forms of *be* are not semantically vacuous, but do contribute some relation, that all modals have unique KEY values, and that *do* bears a KEY value that unifies with the KEY value of any main verb, but no auxiliaries.

[78]Another example of a construction that constrains tense is the set of 'introducers' ***it's time***, ***it's about time*** and ***it's high time***, which generally require past tense in the following indicative clause (Fillmore 1988:51).

(i) It's high time you started thinking about your future.

The third way in which the tag must agree with its sentence is that the subject of the tag must be coindexed with the subject of the sentence. However, as tags are sentence modifiers (as opposed to VP modifiers), information about the subject of the sentence is not directly available through the MOD value. Following Kathol (1999), Bender and Flickinger use the agreement facts to motivate a head feature of nouns and verbs called AGR, that makes this information available at the sentence level.

While most nouns identify their AGR value with their semantic index, the singular *they* series of pronouns do not. The values of AGR and INDEX in Figure 2.52 are designed to account for sentences like (70).

(70) Everyone wins, don't they?

2.6 Summary

This chapter has presented an overview of the framework of Construction-HPSG, with particular attention to analyses that obviate the need for empty categories and the treatment of English auxiliaries. With this as background, we are ready to turn to the analysis of copula absence in AAVE.

Chapter 3

The Syntax of AAVE Copula Absence

3.1 Introduction

This chapter is concerned with the syntax of AAVE sentences like (1), where there is no surface realization of the copula.

(1) You in trouble.

Such copulaless sentences are available in many but not all contexts. At the same time, corresponding sentences with overt forms of the copula seem to be available in all contexts. Thus one of the main points of interest is the relationship between sentences with and without the copula in AAVE.

Throughout the literature on AAVE copula absence, the auxiliary and copula uses of *be* are lumped together. With the exception of Labov's phonological rules discussed in Section 3.3, the term *copula* is generally used to refer to both uses. As discussed in Section 2.5 of Chapter 2, there is good syntactic evidence for treating so-called copula and auxiliary uses of *be* as involving one and the same lexeme. Therefore, I see no reason not to follow this practice here.

Labov (1969, 1995) presents an analysis of AAVE copula absence in terms of phonological deletion of the residue left behind by copula contraction. In Section 3.3,

I review the arguments for this analysis and argue that copula absence is better treated as a syntactic, rather than a phonological, phenomenon.

In Section 3.4, I review and extend the construction-based (i.e., surface-oriented) analysis of copula absence proposed by Mufwene (1992) and Sag and Wasow (1999). This analysis turns out to be falsified by subject extraction facts. Section 3.5 presents a second constructional analysis which also runs into empirical problems. Section 3.6 presents an analysis in terms of a phonologically empty form of *be*. Finally, Section 3.7 presents a third constructional analysis that builds on some of the intuitions of the silent verb analysis, and is able to account for all of the facts considered. Both of these last two analyses can be considered empty category analyses, at some level. Thus this chapter will show AAVE copula absence to be an exception to the general trend away from empty categories reviewed in the last chapter.

First, however, we turn to a brief discussion of data collection, grammaticality judgments, and individual variation in AAVE.

3.2 Data collection

The AAVE data considered in this chapter include some from published sources (principally Labov 1969, 1995), some from the data collected in interviews by the Copula Project at Stanford University, and some invented sentences that have been checked with four native speakers. The sentences from published sources were also checked with the same speakers. The judgments for the invented sentences were consistent across all four consultants unless otherwise indicated.

On top of the usual concerns surrounding grammaticality judgments, there are some additional issues that come up in eliciting grammaticality judgments in stigmatized varieties such as AAVE. First, there is the danger that consultants will deny the existence of any systematicity in the stigmatized variety and claim that "anything goes". However, three of the consultants I was working with are linguists, and the other has taken at least one course on the linguistics of AAVE. They all therefore see AAVE as a systematic variety. Further, they were all willing to say that some of the strings I tested were bad.

The next problem is bidialectalism. The speakers I consulted, like many adult speakers of AAVE, also have command of Standard English. It is therefore possible that certain sentences they judged to be grammatical are grammatical in fact only in the standard variety and not in AAVE. However, the speakers I consulted were willing to point out lexical choices in my sentences that were inauthentic (and correct them), as well as constructions that they believed were only available in the standard variety. It is not possible to tell if they caught everything, of course, but the fact that they pointed out some such inauthentic choices is encouraging.

It is important to note in this connection that copula presence is as much a feature of AAVE as copula absence. Labov (1995:33) illustrates this nicely with "sounding" (ritual insult) data that he collected, including the examples in (2). He writes, "Here the speakers of the language are engaged in intense interaction with each other, using their basic vernacular, and we observe the rapid alternation of zero, contracted and full forms of the copula."

(2) a. Your mother is a Phil D. Basket.
 b. Your mama's a weight lifter.
 c. Your mother a ass, period.
 d. Because he old, he's old, that's why!

Finally, there is no reason to expect AAVE to be any more consistent across speakers than other varieties of English,[1] and my consultants did disagree on some of the sentences. There is always some amount of "noise" in grammaticality judgment data, as grammaticality judgments themselves are an aspect of performance (Schütze 1996). Further there is also legitimate regional variation (and my speakers did come from different parts of the country) as well as variation between individuals within the same community. However, my sample (4 speakers) was too small to tell legitimate variation from noise in the data. All I can do here is report the data in detail. If a sentence was not judged consistently by the different consultants, or if a consultant indicated that it seemed like something other AAVE speakers would say, this will be noted as the sentence is discussed below.

[1] Although it may be more consistent than some regional or national standards that may be learned later in life for most, if not all, speakers.

3.3 The phonological deletion analysis

Labov's (1972, Ch. 3;[2] 1995) deletion account states that copula absence is the result of a phonological rule of deletion fed by copula contraction. That is, AAVE is seen as taking the general English process of copula and auxiliary reduction one step further. This analysis can be seen as embodying two principle claims: (1) that the copula is present underlyingly, even when it is not realized overtly, and (2) that copula absence is only possible in environments where copula contraction is.

Section 3.3.1 reviews the technical details of Labov's analysis. Section 3.3.2 reviews Labov's evidence for the claim that copula absence is only possible where copula contraction is, and provides two robust counterexamples which falsify the claim. Section 3.3.3 reviews the evidence Labov adduces from the variation data for his phonological analysis and Rickford et al.'s (1991) arguments that this evidence is, at best, inconclusive.

3.3.1 Formalization

As formalized in Labov 1972, Ch. 3, this account consists of four SPE (Chomsky and Halle 1968) style rules: The 'weak word rule' reduces the stress of weak ([+W]) words, 'vowel reduction' reduces unstressed lax vowels to schwa, 'auxiliary contraction' deletes the schwa of a tensed vowel-initial auxiliary, and 'auxiliary deletion' deletes the remaining consonant. The rules are formalized as in Figure 3.1

Rule A in Figure 3.1 is designed to interact with Chomsky and Halle's (1968) cyclical nuclear stress rule so that any 'weak word' (in particular, auxiliaries) will have its stress removed completely, if it has already been reduced twice by the cyclical nuclear stress rule.[3] The intended effect is that the auxiliaries should be unstressed, and thus undergo vowel reduction (rule B), just in case they are not in stressed positions such as before ellipsis and extraction sites.

Rules C and D are variable rules, meaning that they apply optionally with some probability and are probabilistically sensitive to certain factors of the context. This is

[2] A revised version of Labov 1969.
[3] I.e., assigned the value '3stress' as opposed to '1stress' or '2stress'.

A Weak word rule

$$\begin{bmatrix} +\text{W} \\ 3\text{stress} \\ \text{V} \end{bmatrix} \longrightarrow \begin{bmatrix} -\ \text{stress} \end{bmatrix}$$

B Vowel reduction

$$\begin{bmatrix} -\text{stress} \\ -\text{tense} \\ \text{V} \end{bmatrix} \longrightarrow \begin{bmatrix} +\ \text{cen} \end{bmatrix}$$

C Auxiliary contraction

$$\begin{bmatrix} +\text{voc} \\ -\text{str} \\ +\text{cen} \end{bmatrix} \longrightarrow \langle \phi \rangle \ / \ \left\langle \begin{matrix} +\text{Pro} \\ -\text{Cons} \end{matrix} \right\rangle \ \#\# \begin{bmatrix} \underline{\quad\quad} \\ +\text{T} \end{bmatrix} \begin{matrix} \text{C}_0^1 \\ \langle *\text{nas} \rangle \end{matrix} \ \#\# \left\langle \begin{matrix} +\text{Vb} \\ +\text{Fut} \\ -\text{NP} \end{matrix} \right\rangle$$

D Auxiliary deletion

$$\begin{bmatrix} +\text{cons} \end{bmatrix} \longrightarrow \langle \phi \rangle \ / \ \left\langle \begin{matrix} *\text{strid} \\ +\text{cons} \\ +\text{Pro} \end{matrix} \right\rangle \ \#\# \begin{bmatrix} \underline{\quad\quad} \\ -\text{nas} \\ +\text{cont} \end{bmatrix} \ \#\# \left\langle \begin{matrix} +\text{Vb} \\ +\text{Fut} \\ -\text{NP} \end{matrix} \right\rangle$$

Figure 3.1: Labov's phonological rules for copula absence

indicated in the notation by angle brackets around the output of the rule (signalling that the rule is a variable rule) and by angle brackets around some of the contextual factors, i.e., those with only probabilistic and not categorical effects. This aspect of Labov's analysis will be discussed briefly in Section 3.3.3 and in more detail in Chapter 5.

The auxiliary contraction rule C applies to an unstressed vowel (schwa), and optionally deletes it just in case it belongs to a tensed (+T) word that consists only of the vowel and possibly one following consonant. (## indicates a word boundary.) In this way, it picks out schwas belonging to tensed auxiliaries. (*Has* contraction would involve an earlier rule of /h/ deletion).

The auxiliary deletion rule deletes consonants that are the sole segment in a word (cf. the ## specifications on either side of the context). Further, it only applies

to non-nasal continuants, thus avoiding deleting /m/ and generating ungrammatical strings like (3).[4]

(3)*I the winner.

The only way to get words that are just one consonant long is to take the output of the auxiliary contraction rule. Thus, contraction feeds deletion. It is important to note in this regard that *contraction* here refers to full contraction, where the auxiliary is reduced to a single consonant. In some contexts, one finds forms that are spelled the same, but in fact the schwa is retained. Labov's deletion rule requires full contraction before it can apply. The next two subsections review Labov's evidence for this analysis and argue that copula absence is better treated as a syntactic phenomenon.

3.3.2 Distributional evidence

Labov's first line of argumentation for the deletion analysis is the claim that copula absence is possible in AAVE just in case copula contraction is possible in Standard English. Labov argues that the two varieties share a basic core grammar, although AAVE has some additional rules, including the rule of auxiliary deletion. To substantiate this claim, Labov shows that copula absence is disallowed in seven contexts that also bar contraction. These are given in Table 3.1. I will discuss each of these environments here as they provide a core set of data that any analysis must account for. Further, in some cases, the data are not as clear as Labov claims, partially because he was working solely with attested data, not grammaticality judgments.[5]

[4] It appears that this constraint might not be as absolute as the literature implies. While all four consultants rejected (3), one accepted (i) and one more recognized (i) as something other speakers might say.

(i) I tired.

[5] "The * symbol is usually used to indicate intuitive reactions of unacceptability: here it refers to patterns of production that are so clear that they support predictions of what is possible or not possible." (1995:39) All of the sentences quoted from Labov in this section have been checked with native speakers. In cases where Labov only illustrated his claims with positive data, I have provided and checked the corresponding ungrammatical examples.

a.	nonfinite contexts
b.	imperatives
c.	ellipsis
d.	emphasis
e.	past tense
f.	inversion
g.	complement extraction

Table 3.1: Environments disallowing both contraction and copula absence

Nonfinite contexts The copula is obligatorily present in non-finite contexts, for example, after *to* in (4).[6]

(4) a. *You got to ϕ good, Rednall!
b. You got to be good, Rednall! (L95)[7]

Further, there is no contracted form of non-finite *be*. On Labov's account, this follows because the form *be*, required in non-finite contexts, is not of the right phonological shape for contraction, so neither contraction nor deletion are possible.

Imperatives The copula is obligatorily present in imperatives, as shown in (5).

(5) a. Be cool, brothers! (L95)
b. *ϕ nice to your mother!

On Labov's account, imperatives are just a subcase of non-finite contexts, as the same form (*be*) is involved in both.

Ellipsis Copula absence is not possible before (verb phrase) ellipsis:

(6) a. (You ain't the best sounder, Eddie!) I ain't! He is! (L95)
b. *They said he wild, and he ϕ. (S&W99)

Contraction is also not possible (in AAVE or other English dialects) in this environment:

[6]In these examples, ϕ indicates the position where the copula is expected but not present.
[7]Key to data sources: L72: Labov 1972, Ch. 3; L95: Labov 1995; S&W99: Sag and Wasow 1999, Ch. 14

(7) a. They said he('s) wild, and he IS.
b. *They said he wild, and he's.

On Labov's account, this follows because *is* is stressed (by virtue of being final in its phrase) and thus can be neither contracted nor, *a fortiori*, deleted.

Emphasis Labov illustrates the emphasis case with the following examples:

(8) a. Allah IS god. (L95)
b. He IS a expert. (L95)

However, if a form of *be* were emphasized and then deleted in some example, the phonological effects of emphasis would be deleted along with the copula Thus counterexamples to this claimed generalization might not be apparent in production data. The only difference between copula absence examples with and without emphasis on the deleted copula is their meaning. The question to ask is whether a sentence like (9), with no special stress on any word, could be used to mean the same thing as (8b).

(9) He a expert.

The answer is appears to be that it cannot.[8]

Inversion Labov includes among his environments which disallow both deletion and contraction two inversion environments: tag questions and matrix polar questions. Tag questions are illustrated in (10).

(10) a. It ain't a flower show, is it?
b. *It ain't a flower show, ϕ it?

[8]I found this out by first asking my consultants directly if (9) and (8b) could mean the same thing. Two said no. For the two who said yes, I followed up by asking whether B's responses in the following short discourses meant the same thing.

(i) A: Don't ask Paul, he's no expert.
B: But he IS a expert.
(ii) A: Don't ask Paul, he's no expert.
B: But he a expert.

Of these two speakers, one detected a contrast, but the other did not.

The situation is somewhat more complicated with matrix polar questions. First, since AAVE, like other English dialects, allows intonation questions like (11), it is not possible to tell where in the string the underlying copula would be in questions like (12).

(11) You're done?

(12) You done?

Negative polarity items would seem to distinguish the two:

(13) a. Are you done yet?
b.*You're done yet?
c. You done yet?

According to this test, the source of (13c) and therefore also (12) would have to be the inverted form. However, (13c) is a grammatical *reduced question* (Hendrick 1982) in Standard English as well. This means that it does not necessarily involve the same mechanisms as other examples of copula absence in AAVE.

It would seem that matrix *wh* questions would provide a better test. Copula absence is allowed in matrix *wh* questions (Sag and Wasow 1999:331):

(14) a. Where he going?
b. Who you talking to?

While (14b) is also acceptable as a reduced question in Standard English, (14a), with a third person singular subject is not (Hendrick 1982).

However, it turns out that the picture isn't much clearer in this case. First, it's not at all clear that AAVE must have the same constraints on reduced questions as Standard English. More importantly, although Sag and Wasow (1999) claim that AAVE, like Standard English, requires subject-auxiliary inversion in matrix *wh*-questions, (15a–b) were accepted by all four consultants and (15c–e) by some.[9]

[9] Curiously enough, none accepted the present tense counterparts of (15d–e), although the contracted forms were rated as marginal (and restricted to certain kinds of speakers, such as the elderly or young children) by two consultants.

(i) *What he is/he's doing?
(ii) *Where he is/he's going?

(15) a. Who you think gon come?
b. What they found there?
c. What he should do?[10]
d. What he was doing?[11]
e. Where he was going?

It appears that for one of these speakers, matrix *wh* questions require inversion just in case the clause is headed by an auxiliary. For the others, matrix *wh* questions allow inversion, but do not require it.[12]

Another inversion context to consider is inversion triggered by *so*. According to Pullum and Zwicky (1997), this is a possible environment for contraction, and my consultants agree that *so's* in (16a) need only be one syllable long. Copula absence, however, is not allowed in this environment:

(16) a. I'm tired and so's my dog.
b. *I'm tired and so my dog.

That this is indeed an inversion context is shown by the ungrammaticality of (17), on the intended interpretation.

(17)*I'm tired and so my dog is.

Perhaps the most general statement of copula absence across these environments is to say that copula absence does not appear where inversion is required, that matrix polar questions such as (12) and (13c) are instances of some other process or construction, and that *wh* questions do not require inversion.

This predicts that the null copula should be barred from other inversion contexts. One such context is exclamatives, such as (18).

[10] (15c) was rated fine by one consultant, fine but like an older person's speech by another, and ? by a third.

[11] (15d) sounded "childish, but not unusual" to one consultant.

[12] It is not immediately clear how to account for the *is* facts in footnote 9. One possibility is to allow constraints on the combination of certain words with certain constructions. Another is that the deviance of these examples is actually an instance of stylistic discord (Sylva and Zwicky 1975): the relatively standard overt *is* may clash with the non-standard uninverted *wh* question. According to Sylva and Zwicky, judgments of stylistic discord are distinguishable from judgments of ungrammaticality, so it may be possible in principle to test whether this is indeed a case of stylistic discord.

(18) Is he ever tall!

It's a little unclear if this is actually a good sentence in AAVE. Two of the speakers consulted accepted it, one rejected it, and one accepted it but identified it as being from another dialect. Nonetheless, all of the consultants rejected the uninverted version (19a) and the copulaless version (19b).

(19) a. *He is ever tall!
b. *ϕ he ever tall!

To the extent that (18) is a sentence of AAVE, these data provide further support for the claim that copula absence is disallowed in inversion contexts.

Note that the facts regarding inversion triggered by *so* complicate Labov's account somewhat in that the examples in (16) require some further (syntactic) restriction on the deletion rule to keep it from applying. Labov was able to reformulate all of the (categorical) syntactic constraints on copula absence that he found in terms of phonological constraints to do with stress placement and/or the identity of the segment deleted (only /r/ and /z/, not /m/). However, since copula absence does not pattern with contraction in embedded inversion contexts, it is not possible to appeal to stress to rule out (16b).[13] This encroachment of syntactic constraints on copula absence can be taken as one piece of evidence that it is perhaps best treated as a syntactic phenomenon. We now turn to some evidence that deletion is not in fact fed by contraction.

Past tense Labov illustrates the claim that all past tense forms of the copula are overt with the positive examples in (20).[14]

(20) a. I was small; I was sump'm about one years o'baby.
b. She was likin me ... she was likin George too.

[13]Note, however, that Pullum and Zwicky (1997) argue that contraction itself is subject to both syntactic and phonological constraints.

[14](20a) was rejected by some of my consultants, for reasons not related to the copula (the plural *years* with the determiner *one* and the form *o'baby*). (20b) was uniformly accepted.

Note that this is a crucial case for Labov: *was* does not contract completely, and examples of copula absence with past tense interpretation would constitute a case of deletion without contraction.

Green (in preparation) cites the following literary example, from John Edgar Wideman's *Brothers and Keepers*, 1984, p. 114.

(21) I dug being militant cause I was good. It was something I could do. Rap to people. Whip a righteous message on em. People knew my name. They'd listen. And I'd steady take care of business. **This was when Rap Brown and Stokeley and Bobby Seale and them ϕ on TV**. I identified with those cats... [emphasis added]

Three of my consultants accept the boldface sentence in this context, and only one of those accepts the same sentence with overt *is/are*. However, the other two both identified the sentence as somewhat surprising, noting the role of context in facilitating its acceptability.

Another possible example comes from a narrative collected by Fay McNair-Knox, although in this case it's difficult to tell whether the copulaless sentences involve the past or the historical present:[15]

(22) They was searchin an stuff, so me an Maria say, "They fiddin ta turn it out, so we fiddin ta go." **So we tryin to find our partners you know, an we hangin, you know**. An then the thing, this white an this Black girl was fightin, huh? They was tearin it off. [emphasis added]

Judging from these attested examples, a more general account of tense and copula absence would say that copulaless sentences are unspecified for tense, and get the common present tense interpretation as a default. However, available syntactic tests would seem to contradict this. In particular, three of the four native speakers consulted rejected the sentences in (23), and the fourth found them somewhat marginal.

(23) a.*He here yesterday.
b.*He here, wasn't he?

[15]This example was checked with two consultants who both accepted the boldface sentence in this context.

It's not clear what exactly is going on with copula absence and tense. It seems that the present tense reading is preferred, to the point where it is the only one apparent in decontextualized examples, but not necessarily the only one possible. Next we turn to two clear counterexamples to Labov's claim.

Complement extraction According to Labov, the copula is obligatorily present if its complement is extracted:

(24) a. I don't care what you are. (L95)
b. *I don't care what you ϕ.

Again, contraction is not possible in this environment in AAVE or other English dialects, and the explanation Labov appeals to is in terms of stress.

(25)*I don't care what you're.

In an attempt to show that it really is stress that's the problem and not just final position, Labov (1995) notes the following data and gives the explanation below.[16]

(26)

AAVE	Other English dialects
a.*Who it?	*Who's it?
b. Who IT? [in a game]	Who's IT?
c.*What it?	*What's it?
d. What it for/Wha's it for?	What's it for?

> We can't say [(26a)] with the dummy *it*, since dummy *it* is not stressable and the stress must be placed on the copula; but we can say [(26b)] with lexical *IT* which accepts stress. We can't say [(26c)], with dummy *it*, since again the copula has to take the stress; but we can say [(26d)], when the word *for* follows to take the main nuclear stress. (1995:40)

Although stress undeniably plays a role in the distribution of contraction, and sentence-final position, equally undeniably, is not the only environment in which the copula receives stress, there's something else going on in the examples in (26). Consider the related sentences in (27).

[16]The string ***what it for*** was accepted by one consultant, rated marginal by one, and identified as something other speakers might say by a third. The fourth consultant rejected it.

(27) a. It's Kim.
b. Kim's IT.
c. It's the mail.
d. It's for cooking.

In (26a) and (26c) (corresponding to (27a) and (27c)), the complement of *is* is extracted. On the other hand, (26b) is a subject *wh* question; the complement of the copula is still in situ. Likewise, in (26d), the extracted element corresponds to something inside of the complement (the complement of *for*), so that the complement of the copula itself is not extracted. In sum, the pattern shown in (26) is just a sub-case of the claimed ban on copula absence in the context of complement extraction.

However, it turns out that the complement of the copula can be extracted, in certain situations:[17]

(28) a. How old your baby?
b. How tall that basketball player?
c. Where your car?
d. When your birthday?
e. I don't know how old his baby.

This is problematic for Labov because, in non-inverted contexts, it leads to minimal sets like (29). Here the putatively deleted form (29b) is possible, but the contracted form (29c) is not.[18]

(29) a. How old you think his baby is?
b. How old you think his baby?
c. *How old you think his baby's

There is actually a second possible 'source' for (29b) for two of the speakers consulted, namely (30a). For one of these speakers, contraction is not possible here, either.

[17]There is some variation across speakers with respect to these sentences. One speaker accepted all of them. One more identified (28d) as something other speakers might say and accepted the rest. The third speaker rated (28b) and (28e) as marginal and accepted the rest. The fourth speaker accepted (28c) and (28e), rated (28a) and (28b) as marginal, and rejected (28d). (28c) was accepted by all four speakers.

[18](29a) was accepted by three of the four consultants. The other judgments in this set are uniform across all four speakers.

(30) a. How old you think is his baby?
b. %How old you think's his baby?

The details of complement extraction with the silent copula will be taken up again in Section 3.6.4.

Subject extraction The minimal set in (31) provides a second example where deletion is possible but contraction isn't:

(31) a. Tha's the man they say is in love.
b. Tha's the man they say ϕ in love.
c.*Tha's the man they say's in love. (with complete contraction)

Sentences (31a), with a full, uncontracted form of the copula, and (31b), with no overt form of the copula, were accepted by all four consultants. (31c), on the other hand, with complete contraction of the copula, is ungrammatical.[19] Not surprisingly, this is a context that appears to block complete contraction in other English dialects as well.[20]

Labov's deletion analysis has some initial plausibility, as it seemed at first glance that there was a significant parallelism to account for between copula absence and contraction. However, we have seen in this section that that parallelism is in fact

[19]One consultant did accept it, but it appears that this was the result of interference of the inflectional *-s*. This consultant does get the contrast once the possibility of that inflection is removed, as in (i):

(i) a. Tha's the man they gon say is in love.
b. Tha's the man they gon say in love.
c.*Tha's the man they gon say's in love. (with complete contraction)

[20]Strangely enough, although a preceding vowel is usually more conducive to contraction (see §3.3.3 below), corresponding sentences with preceding consonants are better. (i) was accepted with complete contraction by three of four consultants.

(i) Tha's the man they said's in love.

Although this fact is puzzling, it does not detract from the force of (31) as a counterexample to Labov's deletion analysis. Labov's analysis makes the strong prediction that contraction must be possible EVERYWHERE deletion is, because deletion is fed by contraction. Any environment, no matter how specifically characterized, where deletion is possible and contraction is not, is a fatal counterexample.

inexact. There is at least one environment (embedded inversion) where contraction is possible without deletion, making it harder to account for the distribution of copula absence in purely phonological terms. More importantly, copula absence is possible and contraction impossible in at least two environments (complement extraction and subject extraction) and possibly one more (past tense contexts). Thus deletion is possible in at least two environments where contraction isn't and copula absence cannot be completely accounted for in terms of a rule that deletes the consonant left by copula contraction. If the deletion rule has to remove the vowel as well, it becomes much less natural as a phonological rule. Finally, as will be shown in Sections 3.6 and 3.7, syntactic analyses can also provide natural accounts of the failure of copula absence when it occurs. Indeed, such an account will be necessary in general, as Ferguson (1971) has shown that other languages (e.g., Russian and Arabic) which allow copulaless sentences and have nothing corresponding to English auxiliary contraction, nonetheless require an overt copula in a set of contexts similar to that found in AAVE.[21] Before proceeding to the syntactic accounts, however, Section 3.3.3 discusses the other type of evidence Labov adduced for his deletion account.

3.3.3 Variable evidence

In addition to the hard constraints on copula absence, Labov studied the soft or probabilistic constraints, by recording naturally occurring speech and observing features of the context of each occurrence of copula absence or copula presence. These constraints will be discussed in more detail in Chapter 4. Here, we will consider the one such constraint that constituted evidence for Labov's argument that copula absence is a phonological phenomenon. Labov found that deletion and contraction differed in the way they were affected by the preceding phonological environment: A preceding vowel favored contraction, while a preceding consonant favored deletion. This is illustrated in Table 3.2, for data from two of the groups Labov studied (adapted from Table 2.2 of Labov 1995:46).[22]

[21]A similar point is made in Rickford et al. 1991.
[22]The rates of deletion and contraction add up to more than 100% because of Labov's method for calculating them, discussed immediately below.

Preceding segment	Rate of contraction	N	Rate of deletion	N
Cobras				
__ C	.41	46	.80	20
__ V	.90	32	.41	29
Jets				
__ C	.32	93	.70	30
__ V	.90	32	.41	29

Table 3.2: Effect of preceding phonological environment, South Harlem

This result would follow nicely from a phonological account. As Labov argues, the favoring effect of a preceding vowel on contraction and the favoring effect of a preceding consonant on deletion can both be seen as reflexes of a pressure towards creating optimal CVC syllables.

However, Rickford et al. (1991) show that this effect is dependent on how one calculates the rates of each variant. In Labov's (1969, 1995) system, since deletion is analyzed as being fed by contraction, both deleted forms and contracted forms were counted as instances of contraction. That is, the rate of contraction is the number of contracted or deleted forms divided by the total number of forms (contracted, deleted, or full). Again, since deletion is fed by contraction, the rate of deletion is just the number of deleted forms divided by the number of contracted forms. These methods of counting forms are known as 'Labov Contraction' and 'Labov Deletion', respectively.

Romaine (1982) proposes an alternative in which zero forms are counted as a proportion of all tokens and contracted forms as a proportion of only the overt tokens. This method of counting contracted forms is known as 'Romaine Contraction'. A third proposal is 'Straight Deletion' (same as Romaine's proposal for deletion) and 'Straight Contraction', where contracted forms are counted as a proportion of all tokens. Rickford et al. (1991) point out that Straight Deletion and Romaine Contraction make more sense in light of the history of AAVE: If AAVE, as a decreolizing variety, has a diachronic process of copula insertion, then there is no reason to look at the zero forms as a proportion of zero + contracted forms. Likewise, there's no reason

to consider the zero forms as potential input to a contraction rule. This amounts to treating AAVE copula absence as a syntactic rather than phonological phenomenon, and is consistent with both of the syntactic analyses presented below.

Rickford et al. (1991) find that (for their East Palo Alto data) the effect of the preceding phonological environment on contraction is constant across the different ways of calculating the proportion of contracted elements: a preceding vowel always favors contraction. However, the effect of the preceding phonological environment on deletion (zero forms) IS dependent on the method of calculation. With Labov Deletion, a preceding consonant favors the zero forms, while a preceding vowel favors the contracted forms. With Straight Deletion, there is no significant effect of the preceding phonological environment. This is illustrated in Table 3.3, excerpted from Rickford et al.'s Table 6 (1991:117). The values given are variable rule factor weights and parentheses around a value indicate that the factor was not chosen as significant by the Varbrul program.[23]

Preceding segment	Labov Deletion	Straight Deletion
__ C	.59	(.47)
__ V	.41	(.53)

Table 3.3: Effect of preceding phonological environment, East Palo Alto

As Rickford et al. note, these results show that the possible phonological effect on copula deletion cannot be used to argue for either a phonological or a syntactic analysis of the phenomenon. If the frequency of each variant is calculated according to Labov's theory that contraction feeds deletion, a phonological effect appears for deletion. If the frequency of forms is calculated according to a theory where the zero form is a separate lexical item, no phonological effect is found.

[23]For more on Varbrul, see Chapter 4.

3.3.4 Summary

In this section we have seen that Labov's phonological analysis of copula absence is not, in fact, supported by the evidence. There are counterexamples to the parallelism between contraction and deletion and the variable evidence is inconclusive. Further, as suggested by Rickford et al. (1991), inasmuch as Labov's analysis embodies historical claims, it is inaccurate. On the deletion account, copula absence is conceptualized as the result of AAVE speakers taking the general English process of contraction one step further. However, in the decades since Labov's original (1969) proposal, much evidence has been found to support the theory that AAVE developed from an earlier creole, rather than by dialect divergence from other English varieties. The original creole is thought to have had categorical copula absence in many environments. This means that the historical process is one of copula insertion rather than copula deletion. For more on the creole origins of AAVE, see Section 5.4.3 of Chapter 5 and Rickford 1998.

3.4 Constructional analysis I

Mufwene (1992) and Sag and Wasow (1999) independently suggest similar constructional analyses of copula absence in AAVE. These constructional analyses avoid the need to posit a rule of deletion or even a phonologically empty verb, and as such, are in keeping with the surface-orientation of much recent work in HPSG.

Building on the observation that copulaless sentences in AAVE have the same distribution as finite clauses (32), the construction posited builds a finite clause out of an NP and a predicative phrase.

(32) a. You ϕ in trouble. (S&W99)
b. If you ϕ alone, watch out! (S&W99)
c. The man she ϕ lookin for ain't here. (S&W99)
d. He just feel like he ϕ gettin cripple up from arthritis. (L95)

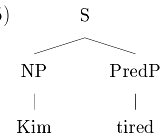
Mufwene's (1992) version is given in (33), where PredP stands for Predicative Phrase, a cover category for NP, AP and PP predicates and participial verb phrases.

(33) S $\rightarrow$ NP PredP

This rule (33) assigns the structure in (35) to the sentence in (34).[24]

(34) Kim ϕ tired.

(35)



Importantly, there is no phrase structural position in this tree for the copula.

3.4.1 Formalization

Figure 3.2 gives the full details the *zero-copula-phrase* proposed by Sag and Wasow (1999), translated into the notation described in Chapter 2. The first thing to notice about this construction is that it is not a headed construction: the syntactic category of the mother is not the same as the syntactic category of either daughter. Copulaless sentences in AAVE have predicates that are not finite verbs, and yet they pattern the same way as regular sentences headed by finite verbs. Accordingly, the *zero-copula-phrase* has no HEAD-DTR feature, but rather two NON-HEAD-DTRS, and it is not a subtype of *headed-phrase*.

The information given in (33) is also encoded in the *zero-copula-phrase*: The mother is a sentence (that is, something finite and verbal with all of its valence requirements satisfied), the first daughter is an NP and the second daughter is a predicative ([PRED +]) phrase.[25] Notice that the value of HEAD on the second daughter is underspecified: It could be anything (*verb, noun, adjective, preposition*), as long as it is also [PRED +]. This is the same natural category selected by *be*.

[24] Accepted by all but one of my consultants, who identified it as something other speakers would say.

[25] Copulaless clauses with non-NP subjects may be possible for some speakers. One consultant accepted (i).

(i) What I want to know who done it?

Any of the analyses discussed here can easily be generalized to allow non-NP subjects.

$$\begin{bmatrix} \textit{zero-copula-phrase} & \\ \text{PHON} & \boxed{A} \oplus \boxed{B} \\ \text{SYNSEM} & \begin{bmatrix} \text{LOCAL} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \textit{verb} & \\ \text{FORM} & \textit{fin} \end{bmatrix} \\ \text{SUBJ} & \langle\ \rangle \\ \text{COMPS} & \langle\ \rangle \end{bmatrix} \\ \text{CONT} & \begin{bmatrix} \text{INDEX} & \boxed{2} \\ \text{RELS} & \boxed{C} \oplus \boxed{D} \end{bmatrix} \end{bmatrix} \end{bmatrix} \\ \text{NON-HEAD-DTRS} & \left\langle \begin{bmatrix} \text{PHON} & \boxed{A} \\ \text{SYNSEM} \ \boxed{1} & \begin{bmatrix} \text{LOCAL} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \textit{noun} & \\ \text{CASE} & \textit{nom} \\ \text{AGR} & \textit{non-1sg} \end{bmatrix} \\ \text{SUBJ} & \langle\ \rangle \\ \text{SPR} & \langle\ \rangle \\ \text{COMPS} & \langle\ \rangle \end{bmatrix} \end{bmatrix} \\ \text{CONT} & \begin{bmatrix} \text{RELS} & \boxed{C} \end{bmatrix} \end{bmatrix} \end{bmatrix}, \begin{bmatrix} \text{PHON} & \boxed{B} \\ \text{SYNSEM} & \begin{bmatrix} \text{LOCAL} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \text{PRED} & + \end{bmatrix} \\ \text{SUBJ} & \langle \boxed{1} \rangle \\ \text{SPR} & \langle\ \rangle \\ \text{COMPS} & \langle\ \rangle \end{bmatrix} \\ \text{CONT} & \begin{bmatrix} \text{INDEX} & \boxed{2} \\ \text{RELS} & \boxed{D} \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix} \right\rangle \end{bmatrix}$$

Figure 3.2: Zero copula phrase

The description of the *zero-copula-phrase* goes beyond the information in (33) in several respects. The noun phrase is required to be nominative and non-1sg,[26] ruling out the sentences in (36).[27]

(36) a.*Him tired.

b.*I tired.

Further, the SYNSEM value of the first daughter is identified with the subject requirement of the second daughter. Recall from Section 2.5 of Chapter 2 that *be* is treated as a raising verb. This entails that its complement is always subject-seeking, even when the complement is a predicative noun phrase. This construction takes those subject-seeking predicative phrases and satisfies that subject requirement directly.

Why not simply generalize the *head-subject-phrase* for AAVE to allow both finite VPs and predicative phrases as the predicate? Doing so would mean that even when the predicate is a finite VP, the so-called *head-subject-phrase* would not be a headed phrase. Further, the *zero-copula-phrase* is not possible with first person singular subjects, and thus is differently constrained from the *head-subject-phrase*.

3.4.2 The basic facts

This constructional analysis provides an immediate account for much of the data noted in Section 3.3.2. In particular, this analysis embodies the claim that copula absence is only possible in contexts where the subject immediately precedes the predicative phrase: It is only possible to make the copulaless sentences by putting the subject together with the predicative phrase, and there are no transformations to move one or the other away. From this, it follows that one will not find copula absence in any of the following environments:[28]

(37) Nonfinite contexts:

*You got to ϕ good, Rednall!

[26]Note that this means that the subhierarchy of agreement types for AAVE is not the same as that given for Standard English in Figure 2.15 of Chapter 2.

[27](36a) was only accepted by one consultant and then with the comment "Only because I know I've heard Southern folk say *Him is tired*." (36b) was accepted by that same speaker and identified by one other as something other speakers might say.

[28]The ϕ indicating the place in the string where the copula is expected is retained in these examples for clarity. However, it should be clear that this analysis posits no such empty element.

(38) Ellipsis:
*They said he('s) wild, and he ϕ.

(39) Tag questions:
*It ain't a flower show, ϕ it?

(40) Inversion triggered by *so*:
*I'm tired and so my dog.

(41) Subjectless imperatives:
*ϕ nice to your mother!

Note that copula absence is licensed in (42), as the subject is immediately followed by the predicate.[29]

(42) a. Where he going?
b. Who you talking to?

Copula absence is not possible with emphasis, because there's nothing there to emphasize:

(43) He a expert.
$\neq$ He IS a expert.

The remaining contexts require somewhat more explanation. Here, I will address imperatives with subjects and the issue of tense. Subject extraction will be discussed in Section 3.4.3 below.

Imperatives with subjects The example in (44) does have the subject immediately followed by the verb, so some other explanation is required for its unacceptability.

(44)*Y'all ϕ nice to your mother!

Recall that the mother of the *zero-copula-phrase* is a finite sentence. If imperatives are built with non-finite forms, the *zero-copula-phrase* will not be usable as an imperative. If, following Ginzburg and Sag (to appear), imperatives are built with finite but semantically distinguished verb forms, it would be sufficient to ensure that the mother here is of the wrong semantic type.

[29]Further, since the *zero-copula-phrase* is not [AUX +], it is suitable to head a matrix *wh* question even in those varieties that require inversion in case the clause is headed by an auxiliary.

Past tense As noted in Section 3.3.2 above, it is not clear whether copulaless clauses are specified as present tense or underspecified with respect to tense, or if different speakers' grammars indeed vary on this point. The constructional analysis is consistent with either conclusion: If copulaless clauses are unspecified with respect to tense, the description in Figure 3.2 will suffice. If copulaless clauses are specified as present tense, the construction itself can carry the present tense semantics. This is achieved by adding to Figure 3.2 the constraints shown in Figure 3.3.[30]

$$\begin{bmatrix} \textit{zero-copula-phrase} & \\ \text{SYNSEM} & \begin{bmatrix} \text{LOCAL} & \begin{bmatrix} \text{CONT} & \begin{bmatrix} \text{INDEX} & \boxed{2} \\ \text{RELS} & \boxed{C} \oplus \boxed{D} \oplus \left\langle \begin{bmatrix} \textit{t-overlap\_rel} & \\ \text{ARG1} & \boxed{2} \\ \text{ARG2} & \textit{now} \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

Figure 3.3: Zero copula phrase — present tense only

Tag questions The fact that tags on copulaless sentences use forms of *be* would seem to pose a problem for this account.

(45) a. They best friends, ain't they?
b. They best friends, aren't they?
c. They not best friends, are they?
d.*They best friends, don't they?
e. They know you, don't they?

However, on the analysis of tags reviewed in Section 2.5.4 of Chapter 2, the choice of tag auxiliary is determined by the KEY value of the main clause. Since the *zero-copula-phrase* is similar in semantic effect to overt forms of the copula, it would not be surprising for it to have the same KEY value, predicting the pattern in (45).

[30]The index $\boxed{2}$ is identified with the index of the predicate daughter and eventually with the INDEX of the main semantic relation contributed by the lexical head of the predicate.

3.4.3 Subject extraction

The preceding subsection showed how much of the basic data on copula absence followed from one entailment of the *zero-copula-phrase* analysis: copulaless clauses are only possible when the subject immediately precedes the predicate. It follows from this, together with a traceless analysis of subject extraction, that the subject of a copulaless clause should not be extractable. Unfortunately, sentences such as (46) are perfectly fine:

(46) a. Tha's the man they say __ ϕ in love.
b. Tha's the man they say __ ϕ a crook.
c. Tha's the man they think __ ϕ famous.

In order to account for these sentences on the present analysis, one would need to allow a subject trace as the first daughter of the *zero-copula-phrase*. This runs counter to the original motivation for this analysis: to avoid positing an empty category.

3.5 Constructional analysis II

One way to avoid the problems raised by subject extraction while maintaining a constructional analysis is to effectively move the construction one node down the tree. The *zero-copula-phrase* proposed in the preceding section built a sentence by combining a noun phrase and a predicate. The new constructional analysis proposed here builds a finite VP out of a predicative phrase. This finite VP can then combine with a subject in the normal way. This is a non-branching construction, used to license trees such as the one in Figure 3.4.

3.5.1 Formalization

The construction that builds the VP in Figure 3.4 is the *predicative-VP* shown in Figure 3.5. It is a non-headed construction with only one daughter. It must be non-headed as the HEAD values of the mother and daughter differ: the mother is a non-predicative finite verb phrase. The daughter is a (not necessarily verbal) predicative phrase. Since the construction has no semantic effect (except perhaps

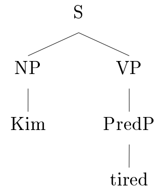


Figure 3.4: Constituent structure under constructional analysis II

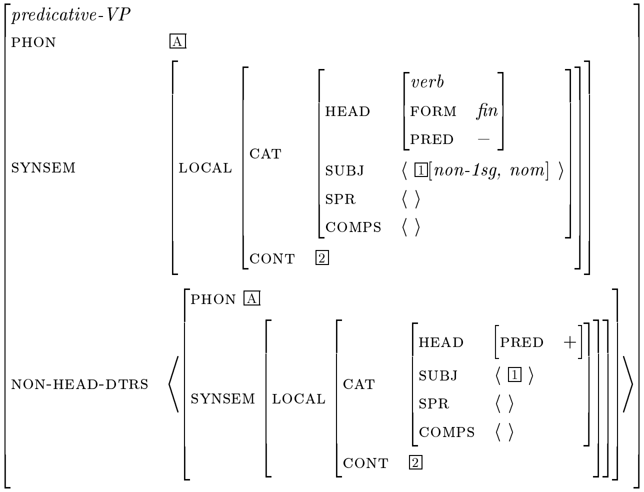


Figure 3.5: Predicative VP construction

to constrain tense, see below), it can simply copy the entire CONTENT value from the daughter to the mother. Similarly, the PHON values are identified. The SUBJ value of the mother and daughter are also identified, but the mother can add further constraints, such as the constraint shown in the figure that the subject be nominative and not first-person singular.

3.5.2 The basic facts

The explanation for the basic facts on this analysis differs somewhat from the previous analysis, but is still fairly straightforward for most cases.

Copula absence is not allowed in nonfinite contexts, as the *predicative-VP* construction is specified as [FORM *fin*].

(47)*You got to ϕ good, Rednall!

Copula absence is not allowed in ellipsis (48), inversion (49), or tag questions (50), as there is no silent auxiliary that could head the *ellipsis-phrase*, *sai-phrase*, or *tag-phrase* (cf. §2.5.3–§2.5.4 of Ch. 2).[31]

(48)*They said he('s) wild, and he ϕ.

(49)*I'm tired and so my dog.

(50)*It ain't a flower show, ϕ it?

Finally, in the remaining three environments, the explanation is the same as it was under the first constructional analysis: The contrast in (51) follows from the fact that there is no auxiliary to emphasize in the copulaless version.

(51) He a expert.
≠ He IS a expert.

Copulaless clauses are disallowed in imperatives because the *predicative-VP* is of the wrong syntactic (FORM) or semantic type.

[31]As before, since the *predicative-VP* is not [AUX +], it would be suitable to head a matrix *wh* question even in those varieties that require inversion just in case the head of the clause is an auxiliary.

(52) a. *ϕ nice to your mother!
b. *Y'all ϕ nice to your mother!

The existence of forms of *be* in tags on copulaless clauses can be modeled by giving the *predicative-VP* the same KEY value as forms of *be*.

(53) a. They not best friends, are they?
b.*They best friends, don't they?

Finally, as with the previous construction, this construction can be specified as present tense or left underspecified for tense.

3.5.3 Further facts: Subject extraction

Because this new constructional analysis does not require the subject to be present in the phrase structure of a copulaless clause, subject extraction (54) can be handled in the usual way.

(54) Tha's the man they think __ ϕ famous.

But what is the usual way? Recall from the review of Bouma et al.'s (in press) analysis in Section 2.3 of Chapter 2 that, in cases of complement extraction, the bottom of an extraction dependency is provided for by the Argument Realization Principle, repeated here as Figure 3.6.

$$\begin{bmatrix} \textit{word} & \\ \text{SYNSEM} & \begin{bmatrix} \text{LOCAL} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{SUBJ} & \boxed{A} \\ \text{COMPS} & \boxed{B} \ominus \textit{list(gap)} \end{bmatrix} \end{bmatrix} \end{bmatrix} \\ \text{ARG-ST} & \boxed{A} \oplus \boxed{B} \end{bmatrix}$$

Figure 3.6: Argument Realization Principle, repeated

The Argument Realization Principle states that the ARG-ST value of a *word* is its SUBJ list, followed by its COMPS list, plus some elements of type *gap* interleaved with the complements. Nothing in this constraint prevents the elements on the SUBJ or COMPS list from also being of type *gap*, however: The value of SUBJ and COMPS is always a list of *synsems*, and *gap* is a subtype of *synsem*. What keeps elements of type *gap* from appearing on COMPS lists is that the resulting words would be unusable: *gap* is a subtype of *non-canonical*, and no *sign* may have a synsem of type *gap*, as all *signs* are constrained to have *canonical* synsems. Since all COMPS members must be realized or otherwise canceled off, a word with *gaps* on its COMPS list could never be used in a grammatical sentence.[32] On the other hand, in the theory of clauses developed by Ginzburg and Sag (to appear), a word with a *gap* on its SUBJ list CAN be used in grammatical sentences. In particular, the complement of proposition embedding verbs like *think* may be [SUBJ ⟨ ⟩] or [SUBJ ⟨ *gap* ⟩].[33]

Therefore, if a word like *famous* could have a *gap* on its SUBJ list, a *predicative-VP* with such a word as its daughter would also have a *gap* on its SUBJ list, and this whole thing could be the complement of *think*. This is sketched in Figure 3.7. The VP whose label is boxed in this figure is the mother of the *predicative-VP*.

[32] The *ellipsis-phrase* and *sai-phrase* allow COMPS elements to unify with the type *elided* and cancel such COMPS requirements off without realizing any corresponding sign in the phrase structure. However, *elided* is a sister type to *gap*, and therefore incompatible with it (cf. Figure 2.50, page 71). Therefore the addition of these phrase types to the grammar does not change the fact that a word with *gaps* on its COMPS list cannot be used in a grammatical sentence. The data are somewhat subtle, but they do support this conclusion. (i) does not involve ellipsis.

(i) Where is there a gas station?

[33] This is not in fact implemented by having verbs like *think* constrain their subject value. Rather, such verbs select for complements with clausal semantics, and a special construction turns VPs specified as [SUBJ *gap*] into clauses, with the appropriate semantic type. For details, see Ginzburg and Sag to appear.

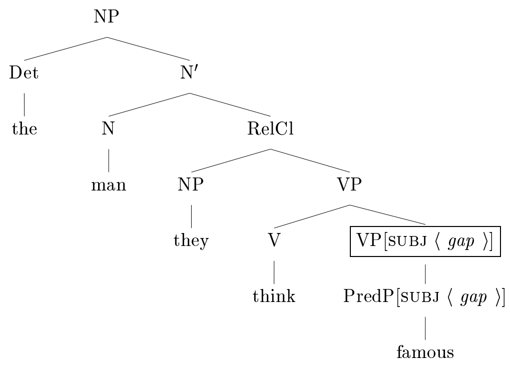


Figure 3.7: Subject extraction with *predicative-VP*

3.5.4 Existentials

This new constructional analysis seems promising: it can account for all of the basic facts plus the subject extraction facts that were problematic for the first constructional analysis. However, it is still insufficient. There are three kinds of facts that it cannot account for, or can account for only with difficulty. The first, to be discussed in this subsection, is the possibility of copulaless existentials. The second, discussed in Section 3.5.5, is a curious interaction of negation and ellipsis. The third, discussed briefly in Section 3.5.6, is the possibility of complement extraction.

The problematic existential sentences are those in (55).

(55) a. There a car blocking my way.
b. I know there at least SOMEBODY happy about this.
c. There a big crowd of people outside.
d. There a book gone from my desk.

It should be noted that these sentences were not unanimously accepted. One speaker rejected all sentences with *there* and a zero copula. The three other speakers all accepted (55b), but each of the other sentences was rated as marginal (although not out) by one speaker, and only accepted by two. Only one speaker accepted all four sentences in (55). It is clear that the only problem for the other three speakers is the form of the copula. With overt *is*, they accept all four sentences.[34]

Nonetheless, it appears that copulaless existentials are a part of AAVE. Despite his claim that *it* is among the class of subjects that always appear with contraction, not deletion, Labov (1969) provides the attested example in (56).

(56) It always somebody tougher than you are.

(*It* is a variant form of the existential expletive in AAVE. (55) gives sentences with *there* because *it* was somewhat marginal for my consultants, especially in copulaless clauses.)

The problem that the sentences in (55) and (56) pose for the *predicative-VP* analysis has to do with their constituent structure. If, as argued below, the constituent

[34] The speaker who rejected all of the sentences in (55) also found the versions of (55d) questionable even with *is*.

structure of existentials is as in Figure 3.8, turning predicative phrases into finite VPs won't help generate these strings. In order for tree (b) of Figure 3.8 to be well formed, the VP *blockin my way* would have to select for two NPs, not the single NP subject that it does select for. Further, there would have to be some construction that realizes both of these NPs before the predicate. The picture is further complicated by sentences like (57), in which the putative head has been extracted:[35]

(57) I know where there a pot of gold.

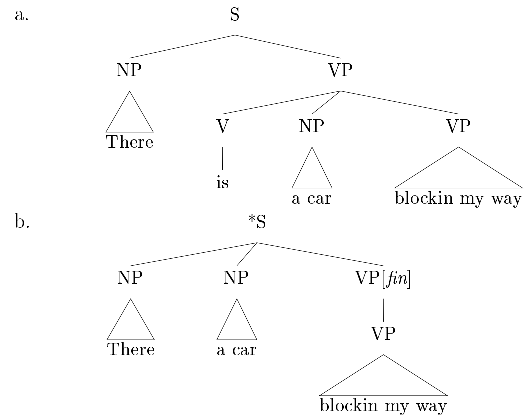


Figure 3.8: Constituent structure of existentials I

One possibility is to account for these sentences by positing instead the constituent structure in Figure 3.9. In this case, the VP *blockin my way* would be a modifier of the NP *a car*, and this whole NP would be the predicative daughter of the *predicative-VP*. This would still be unsatisfactory for a number of reasons. First, it would still require a separate type similar to the *predicative-VP* but distinct in that it selects an expletive subject (*there*). Second, that NP is not predicative: it is the existence

[35] Accepted by two speakers, rated marginal by one, and rejected by the fourth.

of the referent of the NP that is being asserted in such sentences. This semantic intuition is confirmed by the fact that it does not have the syntax of a predicative NP, as illustrated in (58) (cf. (25) on page 48 of Chapter 2).

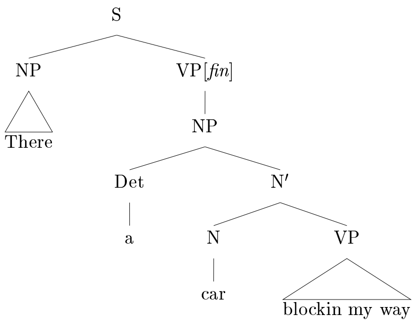


Figure 3.9: Constituent structure of for existentials, II

(58) a.*There was advisor to the committee (in the room).
b.*There advisor to the committee (in the room).
c.*There was too much of a fool to take seriously (in the room).
d.*There too much of a fool to take seriously (in the room).

Third, although this has been the subject of some controversy, there is a variety of evidence that the constituent structure in Figure 3.9 is incorrect: the NP and predicative phrase after the copula in existential sentences are two separate constituents. In the remainder of this subsection, I will review this evidence, principally by refuting each of the arguments that Williams (1984) puts forward for the constituent structure in Figure 3.10. I will call this structure the *be*-NP structure, in contrast to the structure in Figure 3.8 (a), which will be called the *be*-NP-XP structure. These arguments primarily concern the standard varieties studied in the literature cited.

Spurious ambiguity Williams's (1984) first argument is that, given the existence of sentences like (59), one needs to be able to generate existentials with the *be*-NP structure anyway.

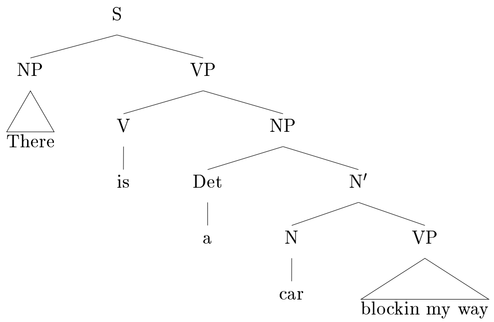


Figure 3.10: *Be*-NP structure for existentials

(59) There is a Santa Claus.

Williams further argues that all existentials could be generated this way. That is, he claims that the material following *be* in an existential is always a possible NP. If this is true, allowing the *be*-NP-XP structure alongside the *be*-NP structure would only lead to spurious ambiguity.

However, it is not the case that the material following *be* in an existential is always a possible NP, if to be a possible NP it must also appear in other NP positions. The following contrast has long been noted (Milsark 1974), and in fact, is noted by Williams (1984:146) himself:

(60) a. There are some people sick.
b. *I saw some people sick.
c. I saw some people sick as dogs.
d. *Some people sick are in the next room.
e. Some people sick as dogs are in the next room.

The generalization is that bare adjectives are not permitted as post-nominal modifiers in English, although they are permitted in existentials. The exception is cases like (61), with an indefinite pronoun:

(61) Someone sick is in the next room.

Further, even when the determiner, noun and following material do make up a possible NP, there are detectable differences in the semantics of this (putative) NP in existential and other sentences. This is illustrated by two sets of examples. In (62),[36] the string *a live fox drowned in a tub of water* is contradictory in (62b) but not in (62a).[37]

(62) a. During the demonstration there was live fox drowned in a tub of water.
b. Jack pulled out a live fox drowned in a tub of water.

To this observation, Lumsden (1988:51) adds the contrasting pairs like (63). The sentence in (63a) does not imply that there are any fish in the tank beyond the oddly colored ones referred to. In contrast, the sentence in (63b) does entail that there are other, less oddly colored fish in the aquarium.

(63) a. There are the most oddly colored fish in that aquarium.
b. The most oddly colored fish in that aquarium swam right at me.

It is plausible that these semantic differences could follow from the difference between the use of the various XPs as predicates and as modifiers. That is, in (63b), the phrase *in that aquarium* is a modifier of *fish*, and falls within the restrictor of the quantifier *most*. In (63a), on the other hand, *in that aquarium* is not part of the NP *the most oddly colored fish*. For whatever reason, this allows the idiomatic sense of *most* which lacks the entailment of a comparison set.

Extraction and the Complex NP Constraint Williams's (1984) somewhat involved second argument is that the impossibility of (64b) is best explained by the restriction on extraction from complex NPs, on a par with (65b).

(64) a. There was someone how happy? [Reprise question]
b. *How happy was there someone?

(65) a. You met someone how happy? [Reprise question]
b. *How happy did you meet someone?

[36]Lumsden (1988:51) attributes this type of example to Milsark (1974:45–46), but no such examples can be found there.
[37]Williams (1984) disputes these judgments, but gives no argument beyond intuitions.

However, this analysis requires him to suppose that examples like (66) are generated by extraposing the modifier *happy/available* from the NP after it is extracted:

(66) a. How many people were there happy?
b. How many unicorns are there available?

Further, extraposition must be constrained to follow extraction, so that it doesn't feed extraction and license sentences like (64b).[38]

In addition to problems of rule ordering, Pollard (1984:123) notes that Williams's (1984) system would require that extraposition is in fact obligatory just in case the postverbal NP of an existential is extracted:

(67)*How many unicorns available are there?

Thus it appears that the ungrammaticality of (64b) is not due to the complex NP constraint. Perhaps an explanation in terms of conflicting pragmatic functions of presentational sentences and *wh*-questions would be more successful. That is, presentational sentences serve to introduce new entities into the discourse. The question in (64b) appears to presuppose the existence of the referent of *someone*.

Heavy NP Shift Williams's (1984) next argument centers on the apparent failure of heavy NP shift in existentials. That is, if there are two constituents following *be* in (68a), one would expect the first to be able to shift over the second, as in the ungrammatical (68b), on the model of (69). Williams points out that the ungrammaticality of (68b) would follow directly if *sick* were in fact part of the NP.

(68) a. There are several of George's recent acquaintances sick.
b. *There are sick several of George's recent acquaintances.

(69) a. I consider several of George's recent acquaintances sick.
b. I consider sick several of George's recent acquaintances.

[38]Williams supposes sentences like (i) are grammatical because PPs can also be sentence modifiers.

(i) I know where there is a unicorn.

However, Quirk et al. (1985) note the grammatical examples in (70) and Thomas Wasow (p.c.) provides the attested examples in (71).

(70) a. There was in the vicinity a helpful doctor.
b. There'll be left open no single well-stocked shop.

(71) a. There was inside his body a feeling of such inevitable danger and ultimate doom that it damn near took his breath away. (John Treadwell Nichols, *The Milagro Beanfield War*, 1974)
b. There must be available no acceptable less discriminatory alternative which would accomplish the purpose as well. (Edward Chen, 'Language Rights in the Private Sector' in James Crawford, *Language Loyalties: A Source Book on the Official English Controversy*, University of Chicago Press, 1992)

Since these examples show that heavy-shift of the post copular NP in existentials is possible, the argument turns out to favor the *be*-NP-XP structure instead. Some other explanation must be found for the ungrammaticality of (68b).

Ellipsis Williams's (1984) argument from ellipsis centers on the apparent impossibility of eliding just the noun in sentences like (72).

(72)*There was someone sick, but there wasn't __ dead.

This contrasts with existentials with PPs after the NP:

(73) There was someone in the parlor, but there wasn't in the garden.

Williams explains this contrast by appealing to the fact that PPs can also appear as sentence modifiers, while APs cannot. However, Warner (2000:205–206) shows that examples such as (74) are only difficult to construct, not impossible. (See also (57–58) on page 69 of Chapter 2.)

(74) Are there any first-year students angry about their grades?
—No, but there are upset with their teachers.

As noted in Section 2.5.3 of Chapter 2, it is not at all clear that examples such as (74) actually involve ellipsis and not pseudo-gapping. Pseudo-gapping famously targets all kinds of strings (Levin 1986), but it does not appear to target initial substrings on NPs:

(75) a. I'm sure I would like him to eat fruit more than I would __ cookies. (Levin 1986:18)
b. *I'm sure I would like him to eat bread made from wheat more than I would __ made from rice.

Thus whether examples like that in (74) involve ellipsis or pseudo-gapping, they undermine Williams's argument.

Non-predicative *with* and cleft existentials Williams's (1984) final argument for the *be*-NP structure is the contrast in (76).

(76) a. The man with a green coat is here.
b. *The man is with a green coat.
c. There is a man with a green coat.

(76a–b) show that PPs headed by *with* may be used only as modifiers and may not be used predicatively—that is, they may not appear as the complement of *be*. Therefore, it is surprising that (76c) doesn't appear to be an 'ontological' existential like (77).

(77) There is a Santa Claus.

Although Williams doesn't bring it up, one could say something similar about examples like (78) with relative clauses:[39]

(78) There's nothing you can do about it.

Note that in (78) the post-copular NP *nothing* is linked to the object of the following (bare) relative clause.

However, as Collins (1992) argues at length, examples such as (78) are best treated as a species of cleft sentences, which differ from standard existentials not only in syntax, but also in pragmatic function. As clefts, they contain two constituents after

[39](78) is an attested example from Collins's (1992) Australian English corpus.

the copula. This raises the possibility that one could get around the problems that the constituent structure of these sentences poses for the *predicative-VP* analysis by positing that copulaless existentials are all standard, and have a *be*-NP structure, while all of the counterexamples to that structure discussed here are in fact clefts. However, this won't do, as the copula can also be absent in cleft existentials:[40]

(79) There always somethin you can try.

As for the *with* examples in (76), perhaps they could be analyzed as clefts. However, even if they can't, they only show that the class of elements in the XP position of *be*-NP-XP is not the same as the class of elements that can follow *be* in non-existentials.

Summary These arguments have shown that existentials are best analyzed with a *be*-NP-XP structure. As argued above, this structure effectively rules out the *predicative-VP* analysis of copulaless existentials. Note that copulaless existentials are also problematic for the *zero-copula-phrase* analysis, as they require a separate (although perhaps related) construction, reducing the parsimony of that analysis. We now turn to another point that is problematic for both of the constructional analyses considered so far.

3.5.5 Ellipsis and *not*

As noted above, copula absence is not possible if the complements of the copula are elided, and this fact receives a natural account on both constructional analyses. However, it turns out that ellipsis is possible just in case it strands *not*, as in the following (attested) example:

(80) They say they're best friends and shit, but they not.

The only way for the constructional analyses to capture this fact is if *not* could be somehow parsed as the predicative phrase, as in Figure 3.11.

[40](79) was accepted by two consultants and identified as something others might say by a third. The speaker who rejected all copulaless *there* clauses also rejected this one.

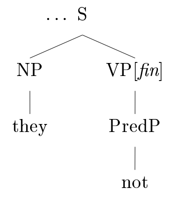


Figure 3.11: Predicative *not*

Not only would this *not* have to be predicative, it would in fact have to license the ellipsis itself, there being no auxiliary available to do so. Given the constructional analysis of ellipsis presented in Section 2.5.3 of Chapter 2, this means that *not* would have to take a predicative phrase as its complement. Such a grammar would produce two parses for all sentences like (81).

(81) They not best friends.

To summarize this brief section, sentences such as (80) can be modeled on either of the constructional analyses discussed so far, but not without the addition of an otherwise unmotivated lexical entry for *not* that would cause considerable spurious ambiguity.

3.5.6 Complement extraction

The second constructional analysis, like the first, predicts that copulaless sentences are only possible when the predicative phrase is in situ. On both analyses, the predicative phrase of a copulaless clause is not the argument of any head. On the traceless analysis of extraction assumed here (see Section 2.3.3 of Chapter 2), this means that it cannot be extracted.

The sentences in (82) show that the complement of the silent copula can, in fact, be extracted.[41] Further, the sentences in (83) show that this is a long distance

[41]There is some variation across speakers with respect to these sentences. See note 17 on page 90 for details.

dependency.[42]

(82) a. How old your baby?
b. How tall that basketball player?
c. Where your car?
d. When your birthday?

(83) a. How old you think his baby?
b. How tall you think that basketball player?
c. How old they say this building?
d. How old they told you his baby?

If these extraction examples could be analyzed a some kind of reduced question (i.e., a separate phenomenon from copula absence), then they could be treated as instances of some other construction. However, reduced questions only affect the matrix auxiliary, as illustrated for Standard English in (84).

(84) a. Who you talking to?
b. *Who do they think you talkin to?

Further, copulaless clauses with the predicative phrase extracted can serve as embedded questions.[43]

(85) a. I don't know how old his baby.
b. I don't care how old they say this building. I still ain't impressed.

Reduced questions, on the other hand, are a matrix clause phenomenon.

Thus the examples in (82) and (83) are bona fide examples of extraction of the predicative phrase in copulaless clauses. They cannot be treated in terms of a separate construction, but require instead a SLASH-based analysis. The only way to do this under the first and second constructional analyses is to revert to a trace-based analysis of extraction. Then a trace could occupy the predicative daughter of either construction and anchor the SLASH dependency. However, doing so would undermine the original motivation of both analyses: to avoid positing empty categories.

[42] Again, there is some variation. (83a) and (83b) were rated ?? by one consultant and accepted by the other three. (83c) was accepted by all four consultants. Three consultants responded for (83d), and they all accepted it, although one said that it would be better with the copula.

Interestingly, for some speakers, (83c) and (83d) appear to require falsetto intonation of the kind studied by Kortenhoven 2000.

[43] (85a) was accepted by all four consultants, although two originally rejected it. (85b) was checked only with the two consultants who originally rejected (85a), and they both accepted it.

3.5.7 Summary

The second constructional analysis improves on the first in that it can handle the subject extraction examples. However, both analyses face problems in existential sentences and sentences with *not* and ellipsis. These facts can be handled, although not elegantly. The final blow is the complement extraction data, which can only be handled at the cost of reintroducing traces. Thus it appears that AAVE copula absence requires some sort of empty category to be posited. In the next section, we turn to an analysis that associates the empty category with the verb itself. This restricts the use of empty categories to a single lexical item (as opposed to traces for all extractions). Further, this silent verb analysis provides more elegant accounts of the existential and *not*/ellipsis examples.

3.6 A silent verb analysis

This section presents an alternative to the constructional analyses: an analysis in terms of a phonologically empty form of the copula. While this is still a syntactic analysis, it does share one of the claims of Labov's deletion account: that copulaless sentences have an underlying copula.

3.6.1 Formalization

On this analysis, the silent form of *be* is modeled as an inflectional type for *be*, on a par with *is*, *are*, etc.[44] The constraints on this type are given in Figure 3.12. As an inflectional type, this type is a subtype of *word*, and has a DAUGHTER feature. In writing the constraints on this type, I am assuming that all of the constraints on the daughter's synsem are carried up to the mother's synsem, except in cases where they conflict.

The PHON value is the empty list, that is, words of this type are phonologically empty. The daughter is constrained to be a *be-lxm*, assuming that all lexemes of the

[44] That is, it's a subtype of *complex-lexeme*, cf. the discussion of morphology in §2.5.1 of Chapter 2.

be family (including the ordinary copula and the *there* copula) are subtypes of this type.[45]

$$\begin{bmatrix} \textit{silent-be} & & \\ \text{PHON} & \langle\,\rangle & \\ \text{SYNSEM} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \text{FORM} & \textit{fin} \\ \text{INV} & - \end{bmatrix} \\ \text{COMPS} & \textit{nelist}(\textit{canonical}) \\ \text{ARG-ST} & \langle\ [\textit{non-1sg}],\ \ldots\ \rangle \end{bmatrix} \end{bmatrix} & \\ \text{DAUGHTER} & \textit{be-lxm} & \end{bmatrix}$$

Figure 3.12: Silent *be* inflectional type

3.6.2 The basic facts

Somewhat surprisingly, the few constraints in Figure 3.12 are sufficient to account for all of the basic facts, as well as most of the facts that were troublesome for the constructional analyses. The exception is the complement extraction facts, the discussion of which will be postponed until Section 3.6.4.

The failure of copula absence with first person singular subjects (86) follows from the agreement constraint that *silent-be* places on its subject. Although this is done in essentially the same way as in the previous two analyses, it is especially natural here: one of the main functions of verbal inflectional types is to place agreement constraints on the subject.

[45] Actually, this may need to be constrained further, as certain other uses of *be* appear to not take the silent form, including *be* in pseudoclefts (i) and other complementation patterns such as (ii).

(i) *A good long sleep what you need most.
(ii) *It because of the cold weather.

((i) and (ii) were rejected by three of the four consultants).
The identity copula, on the other hand, can take the silent form:

(iii) A: Ain't Lew Alcindor Mohammed Ali?
B: You wrong. CASSIUS CLAY Mohammed Ali.

Presumably, the type hierarchy of *be* lexemes can be arranged appropriately.

(86)*I the winner.

The constraint [FORM *fin*] will suffice to keep silent forms of *be* out of non-finite contexts:

(87)*You got to ϕ good, Rednall!

As before, this may account for the lack of copula absence in imperatives as well. If not, an appropriate semantic constraint can be added to the inflectional type.

(88) a. *ϕ nice to your mother!
b. *Y'all ϕ nice to your mother!

The constraint on the COMPS list accounts[46] for the lack of copula absence with ellipsis, tag questions, and inversion triggered by *so*. This constraint says that the COMPS list may not be empty, and it may not have any *non-canonical* synsem types on it.

Recall that on the analysis of ellipsis and tag questions, elements on the COMPS list are unified with the *non-canonical* synsem type *elided*. As this contradicts the constraint on the COMPS list of *silent-be*, (89) and (90) are out.

(89)*They said he('s) wild, and he ϕ.

(90)*It ain't a flower show, ϕ it?

It's not clear if the example in (91) involves ellipsis or an ARG-ST/COMPS mismatch, but either way it would violate the constraint on COMPS.

(91)*I'm tired and so my dog.

The constraint on the COMPS thus accounts for the lack of copula absence in two inverted contexts (tag questions and after *so*.) However, inasmuch as sentences like (92a) are acceptable in AAVE, the silent copula would have to be barred from inverted contexts to rule out (92b). This is achieved by the constraint [INV −] in Figure 3.12.

(92) a. Is he ever tall!
b. *ϕ he ever tall!

[46] Suggested to me by Ivan Sag (p.c.)

Forms of *be* in tag questions on copulaless sentences are unremarkable, as the copulaless sentences are headed by a form of *be*.

(93) They not best friends, are they?

As before, the tense facts can be accommodated by either specifying present tense (specifying a tense value being another well-known function of inflectional types), or leaving tense underspecified, according to how the data actually turn out.

Turning now to the last of the basic facts, something extra will need to be said about the fact that *silent-be* can't be emphasized. However, it's pretty clear what that something extra will be: a general (and natural) constraint against emphasizing phonologically empty words.

(94) He a expert.
≠ He IS a expert.

3.6.3 Further facts

The silent verb analysis can also elegantly account for the further facts that were problematic for the constructional analyses.

First, subject extraction proceeds as with any verb, as this analysis posits a verb in the phrase structure.

(95) Tha's the man they think famous.

Second, *silent-be* is an inflectional type that can take different *be* lexemes as its DAUGHTER value, just like the inflectional types that produce *is*, *are*, etc. In particular, it is a possible inflection of the *there* copula, producing sentences such as (96).

(96) There a car blockin my way.

Finally, as an [AUX +] word, forms of *be* inflected with the *silent-be* type can unify as the DAUGHTER value of *not-arg* (Figure 2.42, page 61), the type that adds

sentential *not* to the ARG-ST list of an auxiliary.[47] The presence of this *not* intuitively satisfies the constraint on the COMPS list of the auxiliary:

(97) They say they best friends and shit, but they not.

Technically, it is a little less straightforward. There are two ways to make sure that (97) is licensed. The first is to have the type *not-arg* not carry up the constraint on the COMPS list. The second is to change the constraint on the comps list to the one in Figure 3.13. This new constraint merely requires that some (initial) element of the COMPS list be canonical. In this case, *not* in (97) satisfies this constraint, and the second complement of *be* is free to unify with the type *elided*.

$$\begin{bmatrix} \textit{silent-be} \\ \text{SYNSEM} \mid \text{CAT} \mid \text{LOCAL} \mid \text{COMPS} \quad \textit{nelist}(\textit{canonical}) \oplus \textit{list} \end{bmatrix}$$

Figure 3.13: A further constraint on the silent *be* inflectional type

Note, however, that an intricate interaction of the facts presented so far actually poses somewhat of a problem for the silent verb account. Recall from Section 3.3 that, for one of the consultants in this study, matrix *wh* questions require inversion just in case the head is an auxiliary. The sentence in (97) shows that the silent verb must be an auxiliary, for only auxiliaries can license ellipsis. However, as shown in Section 3.6.2 above, *silent-be* must be barred from inversion contexts. Therefore, this analysis predicts that (98) should be ungrammatical for the speaker who requires auxiliaries to invert in *wh* questions, and they're not.

[47]It has to be excluded from unifying with the DAUGHTER of *n't-form*, given the ungrammaticality of (i), accepted by only one speaker. (This speaker added that it was good as a reduced form of *isn't*, but not as a reduced form of *ain't*.)

(i)*Bob ϕn't here.

This could be accomplished by leaving the value of $f_{n't}$ undefined for the empty string.
Surprisingly, (ii) and (iii) were accepted by all and all but one of the consultants, respectively.

(ii) They n't here.
(iii) He n't here.

Given the contrast with (i), it appears likely that *theyn't* and *hen't* are contracted forms of *they* + *ain't* and *he* + *ain't*.

(98) a. Where he goin?
b. Who you talkin to?

This problem may find its resolution once the properties of reduced questions in AAVE are properly understood.

3.6.4 Complement extraction

The constraint on the COMPS list in Figure 3.12 is meant as an encoding of the generalization that the silent verb requires at least one of its complements to be canonically expressed. However, as shown in Section 3.3.2 and 3.5.6, this generalization does not hold. The examples in (99) and (100) involve extraction of the (sole) complement of the silent copula—the same silent copula that supports ellipsis only when there is some complement left over.

(99) Where your car?

(100) How old you think his baby?

It might seem possible to save the generalization that the silent copula must appear with at least one complement by reverting to a subject-as-complement analysis of inversion (see, e.g., Sag and Wasow 1999). For example, in (99) *your car*, although notionally the subject of the silent copula, would appear on its COMPS list instead. However, as argued in Section 3.3.2 above the silent copula is systematically barred from inversion contexts. Analyzing the subject as a complement in (99) and (100) would predict that the sentences in (101) are also grammatical.

(101) a. *I'm tired and so my dog.
b. *It ain't a flower show, it?

Further, for some speakers at least, the embedding verb *care* in (102) doesn't take an inverted complement (103):[48]

(102) I don't care how old they say this building. I still ain't impressed.

[48](102) and (103) were only checked with two consultants. Both accepted (102). One rejected (103), and the other accepted it, noting that it sounds better if *say is* is more like *says*. Thus this could be interference from optional *-s* inflection on the verb.

(103) %I don't care how old they say is this building. I still ain't impressed.

We have seen that reverting to a subject-as-complement analysis of inversion won't do the trick. It is possible, however, to account for all of the facts by changing the constraint on the COMPS list. The new constraint needs to capture the facts that the COMPS list may be empty, and it may have some elements of type *elided*, but in the latter case it must also have some elements of type *canonical*. The constraint shown in Figure 3.14 expresses just that.

$$\begin{bmatrix} \textit{silent-be} \\ \text{SYNSEM} \mid \text{CAT} \mid \text{LOCAL} \mid \text{COMPS} \quad \textit{elist} \vee (\textit{nelist}(\textit{canonical}) \oplus \textit{list}(\textit{elided})) \end{bmatrix}$$

Figure 3.14: Revised constraint on the COMPS list of silent *be*

On this analysis, (104) is bad because the COMPS list consists of one element of type *elided*. It thus satisfies neither possible value of the COMPS list: it is not an empty list (*elist*), nor is it a list which begins with some (non-zero) number of *canonical* elements.

(104) *They say he('s) wild and he ϕ.

The same line of reasoning applies to the tag examples like (105), on the ellipsis-based analysis of tag questions given in Section 2.5.4 of Chapter 2. In fact, tags are doubly ruled out, as the silent copula is barred from inversion contexts.[49]

(105) *It ain't a flower show, ϕ it?

(106), on the other hand, is good. In this case the COMPS list consists of one *canonical* element (*not*) followed by one *elided* element, and satisfies the second possible value of the COMPS list.

(106) They say they're best friends and shit, but they not.

[49]The lack of the silent copula in inversion contexts also accounts for the ungrammaticality of (i), even if it doesn't involve ellipsis.

(i) *I'm tired and so my dog.

Likewise, examples like (107) are good because the COMPS list is empty (satisfying the description *elist*). This is so because the Argument Realization Principle provides for *gaps* on the ARG-ST to be absent from the COMPS list (cf. page 38 of Chapter 2).

(107) Where your car?

Something more does need to be said to rule out the examples in (108).

(108) a. *I don't care what you.
b. *Who it?
c. *What it?

A first pass would be a constraint on *silent-be* that any complement *gaps* must not be [HEAD *noun*]. However, some speakers accept the sentences in (109).[50]

(109) a. Who he?
b. Who they?

I leave the investigation of the constraints on the type of complements which can be extracted to future research.[51]

3.6.5 Summary

This section has developed the silent verb analysis and shown that it can account for a wide array of facts. The next section develops a constructional analysis that encodes many of the same ideas. Indeed, the principle difference between the silent verb analysis and the third constructional analysis is the way they interact with the auxiliary constructions.

[50](109a) was accepted by two speakers and reject by the other two. (109b) was accepted by one speaker, rated marginal (? or ??) by two others, and rejected by the fourth.

[51]A further complexity in this vein is the ungrammaticality of (ia):

(i) a. *When they told you the meetin?
b. When they told you the meetin is?
c. Where they told you the meetin at?

(One consultant rejected in (ib) with *told you*, but showed a similar pattern of contrast with *did they tell you* in all of the examples instead.)

As shown in (83d), complements of the silent copula can be extracted over *they told you*. (82d) shows that *when* can be extracted from the complement position of the silent copula. Thus it is not clear why extracting *when* over *they told you* should be bad.

3.7 Constructional analysis III

This third constructional analysis is a formalization of the intuition that in copulaless clauses, the valence of *be* is orchestrating the sentence, without *be* itself being present. In this way it is similar to the silent verb analysis (and Labov's deletion analysis) in positing, in some sense, an underlying copula.

3.7.1 Formalization

Technically, this analysis involves a slight reconceptualization of the daughters features.[52] Under this reconceptualization, there is no NON-HEAD-DTRS feature, but rather a DTRS feature listing all of the daughters. This list of daughters represents the words or phrases that are parsed by the construction. The HEAD-DTR feature then becomes a pointer to one member of this list. Under this reconceptualization, the *head-complements-phrase* is as in Figure 3.16. This is identical to the original version of the *head-complements-phrase*, except that the NON-HEAD-DTRS feature is replaced by the DTRS feature and the value of HEAD-DTR is identified with the first member of the DTRS list ($\boxed{1}$).

To simplify the presentation, I will assume that the *head-complements-phrase* inherits the constraints in Figure 3.15 from a supertype, such as *headed-phrase*. The version of *headed-phrase* in Figure 3.15 is somewhat simpler than the earlier version, since all of the daughters are now in one list. For example, the new version can identify the PHON value of the mother with the append of the PHON values of the daughters. Different subtypes will identify their HEAD-DTR value with different positions in the DTRS list. This *headed-phrase* also allows its subtypes to have constructionally introduced relations. The CX-RELS feature stores the semantic contribution of the construction, which is appended with the relations from the daughters to give the RELS feature of the mother.[53]

[52]This reconceptualization is due to Christopher Callison-Burch and inspired by the LinGO grammar's use of an ARGS feature listing all of the daughters, in addition to the usual HEAD-DTR and NON-HEAD-DTRS features. Its applicability was pointed out to be Ivan Sag, p.c.

[53]This is, of course, possible with the earlier *headed-phrase* as well. I've only introduced it just now because it is about to become relevant.

$$\begin{bmatrix} \textit{headed-phrase} & \\ \text{PHON} & \boxed{C} \oplus \ldots \oplus \boxed{N} \\ \text{SYNSEM} & \begin{bmatrix}\text{LOCAL} & \begin{bmatrix} \text{CAT} & \begin{bmatrix}\text{HEAD} & \boxed{2}\\ \text{SPR} & \boxed{A}\end{bmatrix} \\ \text{CONT} & \begin{bmatrix}\text{INDEX} & \boxed{3} \\ \text{RELS} & \boxed{B}\oplus\boxed{M}\oplus\ldots\oplus\boxed{Z} \\ \text{KEY} & \boxed{4}\end{bmatrix}\end{bmatrix}\end{bmatrix} \\ \text{CX-RELS} & \boxed{B} \\ \text{HEAD-DTR} & \begin{bmatrix}\text{SYNSEM} & \begin{bmatrix}\text{LOCAL} & \begin{bmatrix}\text{CAT} & \begin{bmatrix}\text{HEAD}&\boxed{2}\\ \text{SPR}&\boxed{A}\end{bmatrix}\\ \text{CONT} & \begin{bmatrix}\text{INDEX}&\boxed{3}\\ \text{KEY}&\boxed{4}\end{bmatrix}\end{bmatrix}\end{bmatrix}\end{bmatrix} \\ \text{DTRS} & \left\langle \begin{bmatrix}\text{PHON} & \boxed{C}\\ \ldots\ \text{RELS} & \boxed{M}\end{bmatrix} \ldots \begin{bmatrix}\text{PHON}&\boxed{N}\\ \ldots\ \text{RELS} & \boxed{Z}\end{bmatrix}\right\rangle \end{bmatrix}$$

Figure 3.15: Headed-phrase (revised)

$$\begin{bmatrix} \textit{head-comps-ph} & \\ \text{SYNSEM} & \begin{bmatrix}\text{LOCAL} & \begin{bmatrix}\text{CAT} & \begin{bmatrix}\text{SUBJ} & \boxed{B} \\ \text{COMPS} & \langle\ \rangle\end{bmatrix}\end{bmatrix}\end{bmatrix} \\ \text{HEAD-DTR} & \boxed{1}\begin{bmatrix}\text{SYNSEM} & \begin{bmatrix}\text{LOCAL} & \begin{bmatrix}\text{CAT} & \begin{bmatrix}\text{SUBJ} & \boxed{B} \\ \text{COMPS} & \mathbf{sts}(\boxed{A})\end{bmatrix}\end{bmatrix}\end{bmatrix}\end{bmatrix} \\ \text{DTRS} & \langle \boxed{1} \rangle \oplus \boxed{A} \end{bmatrix}$$

Figure 3.16: Head-complements phrase (revised)

The *silent-copula-phrase* (Figure 3.17), another subtype of *headed-phrase*, is similar to the new *head-complements-phrase*, except that its head daughter is constrained to be a *be-word*,[54] and the construction fails to identify its head daughter with any element of the DTRS list.[55] It also differs in the relationship it posits between the COMPS value of the head daughter and the other daughters. This constraint is written so as to allow complement extraction and partial ellipsis, but ellipsis of all of the complements (for details, see below). Finally, in order to incorporate the relations introduced by the (unexpressed) head daughter, the *silent-copula-phrase* identifies its CX-RELS value with the RELS value of the head daughter ($\boxed{C}$).

$$\begin{bmatrix} \textit{silent-copula-phrase} & \\ \text{SYNSEM} & \begin{bmatrix}\text{LOCAL} & \begin{bmatrix}\text{CAT} & \begin{bmatrix}\text{SUBJ} & \boxed{B} \\ \text{COMPS} & \langle\ \rangle\end{bmatrix}\end{bmatrix}\end{bmatrix} \\ \text{CX-RELS} & \boxed{C} \\ \text{HEAD-DTR} & \begin{bmatrix} \textit{be-word} & \\ \text{SYNSEM} & \begin{bmatrix}\text{LOCAL} & \begin{bmatrix}\text{CAT} & \begin{bmatrix}\text{HEAD} & \begin{bmatrix}\text{FORM} & \textit{fin}\end{bmatrix} \\ \text{SUBJ} & \boxed{B}\langle\ [\textit{non-1sg}]\ \rangle \\ \text{COMPS} & \mathbf{sts}(\boxed{A}\textit{elist}) \vee \\ & (\mathbf{sts}(\boxed{A}\textit{nelist}) \oplus \textit{list}(\textit{elided}))\end{bmatrix} \\ \text{CONT} & \begin{bmatrix}\text{RELS} & \boxed{C}\end{bmatrix}\end{bmatrix}\end{bmatrix}\end{bmatrix} \\ \text{DTRS} & \boxed{A} \end{bmatrix}$$

Figure 3.17: Silent copula phrase

[54]Where *be-word* is assumed to be a supertype to all of the *be* inflectional types.

[55]A similar effect could be achieved under the normal interpretation of HD-DTR and NON-HD-DTRS by positing a construction that was unusual in not incorporating the phonology of its head daughter. However, in that case, one would wonder why there wasn't a similar construction that suppressed the phonology of an arbitrary non-head daughter.

3.7.2 The basic facts

The agreement restriction on the subject is handled by the by-now familiar agreement constraint, encoded by this construction as with the previous constructions.

(110) *I the winner.

Similarly, the lack of copula absence in non-finite contexts is ruled out (as usual) by the [FORM *fin*] specification.

(111)*You got to ϕ good, Rednall!

Again, the failure of copula absence in imperatives is explained either by the FORM value, or by a semantic incompatibility.

(112)*ϕ nice to your mother!

Tags with forms of *be* are again unsurprising, as *silent-copula-phrase*s are headed by a form of *be*.

(113) They not best friends, are they?

The examples involving missing complements are handled in much the same way as in the silent verb analysis. (114) is out because the COMPS value of the head daughter would be the list containing a single *elided* element. This doesn't meet either half of the disjunctive constraint in Figure 3.17: it is neither an empty list, nor a non-empty list beginning with a *canonical* element.

(114)*They say he('s) wild and he ϕ.

In (115), the COMPS value of the head of the *silent-copula-phrase* would be a list containing the (*canonical*) *synsem* of *not* and a single *elided* element. This satisfies the second half of the disjunctive constraint.

(115) They say they're best friends and shit, but they not.

Finally, the complement extraction examples (e.g., (116)) are allowed because they result in an empty COMPS list, satisfying the first half of the disjunctive constraint.

(116) Where your car?

In the remaining contexts, however, the explanations begin to differ. Since copulaless clauses are built by a construction (and not a silent verb), there is once again no auxiliary to head a tag question or an inverted clause. (There is also no subtype of the *silent-copula-phrase* that shares the properties of the *tag-phrase* or *sai-phrase* constructions.)

(117)*It ain't a flower show, ϕ it?

(118)*I'm tired and so my dog.

As for the past tense examples, in this case there is no way to leave copulaless sentences underspecified for tense. The construction can allow either past or present forms of *be* as its head, giving both past and present tense examples, or restrict its head to being present tense only.

3.7.3 Further facts

Subject extraction examples like (119) also pose no problem. Versions of ordinary *be* with a *gap* on their SUBJ list can unify with the *silent-copula-phrase*, and the SLASH value will get passed up in the ordinary way.

(119) Tha's the man they say in love.

Finally, copulaless existentials are also permitted, as the value of the HEAD-DAUGHTER of the construction is defined so as to unify with any *be* lexeme.[56]

(120) There a car blocking my way.

[56]In fact, as noted in footnote 45 on page 119, this may be too general: some *be* lexemes do not participate in this construction. However, given the hierarchical organization of the lexicon, it should not be a problem to find an appropriate type to use to constrain the value of HEAD-DTR in the *silent-copula-phrase*.

3.7.4 Summary

The *silent-copula-phrase* is unusual for a constructional analysis in that the construction will have no phonological content in certain instances. In particular, the complement-extraction example (116) is assigned the structure in Figure 3.18. The boxed node is the mother of the *silent-copula-phrase*. It has no daughters, as the head daughter (*be*) is not included in the DTRS list, and the complement is extracted, but a traceless analysis of extraction is assumed. Thus it is not clear if this constructional analysis really differs from an empty category analysis.

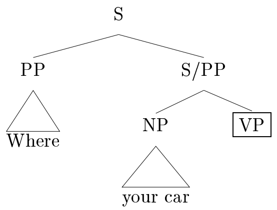


Figure 3.18: Instance of *silent-copula-phrase* with no daughters

3.8 Conclusion

This chapter has reviewed five possible analysis of copula absence in AAVE: a phonological deletion analysis, a silent verb analysis, and three distinct constructional analyses. Each of the first three analyses highlighted new data on copula absence in AAVE. These data show that AAVE copula absence is a syntactic, not phonological, phenomenon. They further show that AAVE copula absence constitutes a case that requires an empty category of some sort in its analysis: either a phonologically empty verb, or a potentially phonologically empty construction. This is an interesting result, given the general success of lexicalist grammars in avoiding empty categories. Nonetheless, I believe that the presence of one necessary empty category does not entail that empty categories should be available as a general tool in

syntax. The psycholinguistic evidence cited in the motivation of surface-oriented, performance-plausible grammars (see e.g., Sag 1992) is not invalidated by one linguistic phenomenon that requires an empty category. Rather, it militates against the proliferation of empty categories, for surface-orientation is a matter of degree. The further grammars are removed from surface strings, the more complicated it becomes to embed them in theories of performance.

With this view of grammar in general and of copula absence in particular as background, we now turn to the issues raised by the use of copula absence in AAVE as a sociolinguistic variable.

Chapter 4

Non-categorical constraints in perception

4.1 Introduction

The previous chapter investigated the hard (or categorical) constraints on AAVE copula absence. AAVE copula absence is also subject to soft (or non-categorical) constraints. As noted, copulaless sentences like (1) almost always have counterparts with full (2) or contracted (3) forms of the copula.

(1) She my piano teacher.

(2) She is my piano teacher.

(3) She's my piano teacher.

The distribution of these three variants is systematically constrained by grammatical factors. These systematic yet violable constraints are well studied in the variationist tradition of sociolinguistics. However, with some important exceptions (see Chapter 5), they are not generally countenanced or even considered by competence models. Of course, it is not obvious that such constraints should be within the purview of competence grammar. The patterns found in production studies could well be the result of systematic performance factors. Therefore, this chapter presents an experiment

designed to show that speakers have knowledge of a particular non-categorical constraint on copula absence. The issue of whether this knowledge is to be distinguished from knowledge of grammar is taken up in Chapter 6.

The structure of this chapter is as follows. Section 4.2 briefly reviews the notion of a sociolinguistic variable, its relationship to other linguistic systems, and the body of work on AAVE copula absence as a sociolinguistic variable. Section 4.3 presents the hypothesis to be tested in the experiment. Section 4.4 describes the methodology employed. Section 4.5 presents the results.

4.2 Non-categorical constraints in production

This section provides some background on AAVE copula absence as a sociolinguistic variable. The first question is why the alternation between overt and zero forms of the copula should be seen as a linguistic unit of any sort. To answer this, Section 4.2.1 reviews some definitions of linguistic variables. Section 4.2.2 applies those definitions to AAVE copula absence. Sections 4.2.3 and 4.2.4 discuss the properties that make the alternation a sociolinguistic variable: its correlations with social and grammatical factors.

4.2.1 Linguistic variables

This section reviews the notion of a linguistic variable with an eye towards whether variables constitute units that should be represented in a competence grammar. Many phonological variables that have been studied involve alternations between allophones of the same phoneme.[1] Thus, to the extent that the phoneme constitutes a unit of grammar, the variable does as well.

However, variationist studies have not been limited to alternations between allophones. In the words of Hudson (1996:169) "[a] variable is a collection of alternatives which have something in common." This is a notion that is applicable at pretty much

[1]Variables such as (r)—the alternation between [ɹ] and [ϕ]—are possibly better conceptualized as the realization or non realization of a phoneme, depending on whether the [ϕ] variant also involves some effect on the preceding vowel.

any level of linguistic structure. Wolfram (1991:23) enumerates six typical ways in which variants are seen to have something in common:

- Structural category realization: The variants are the realization and the non-realization of e.g., a morpheme or phoneme.
- Allo-forms within a structural category: The variants are allophones or allomorphs.
- Processes affecting linguistic classes of units: The variants are the application or non-application of linguistic 'processes', such as word-final consonant-cluster reduction.
- Co-occurrence relationships: The variants are the 'obeying' or 'disobeying' of co-occurrence restrictions, as in variable negative concord or variable subject-verb agreement.
- Item permutation: The variants are different possible orders of linguistic elements as in variable metathesis or variable word order.
- Lexical choice: The variants are (near) synonyms, usually function words.

Note, in regards to the last item, that Lavandera (1978) and Romaine (1981) argue that in variables where the variants bear 'referential' or 'ideational' meaning, one never finds exact equivalence—the variants are different ways of saying similar things.

With respect to representation in the grammar, these types of relatedness form a heterogeneous set. In the case of "process" variables, the process itself would be a thing in the grammar. In the case of lexical choice variables, one wouldn't want to say that the choice between the lexical items is a thing in the grammar. Rather, it is inherent in the grammar, i.e., inherent in the similarities between the two lexical items that allow them to substitute for one another. Thus, depending on the type of variable, the alternation may or may not correspond to something in the grammar. Each variant (e.g., lexical item, allophone, etc.), of course, is something in the grammar.

There are also methodological issues that remove variables somewhat from the domain of individual competence. As Wolfram and Hudson both point out, the identification of variables and of instances of the variants in a corpus is never completely unproblematic. For example, Wolfram's formulation of the co-occurence relationships type of variable makes it clear that, in order to conceive of such a variable, one has to establish the co-occurence restrictions on the basis of something else, perhaps a related variety without the variation.

Hudson gives the example of variable realization of /h/ word-initially (an example of structural category realization). In all varieties of English, there are words spelled with *h* which are never pronounced with *h* (e.g., *hour* in standard English). But why should the membership of this word class be assumed to be the same for all speakers? That is, if a speaker of a different variety of English is heard to say *house* as [aus], one cannot tell on the basis of that one instance whether or not *house* belongs to the *hour* class FOR THAT SPEAKER. Since it is not usually feasible to determine for every speaker studied which words have an underlying /h/, some compromise must be made. Most commonly, according to Hudson, researchers rely on their own knowledge of some related variety (usually a more standard variety).

It is this kind of compromise that distances the linguistic variable from studies of individuals' linguistic competence. Indeed, Wolfram writes:

> Is the cognitive basis of language patterning really one and the same with the sociopsychological patterning that determines the social distribution of linguistic variables? ... It may just be that the patterning of linguistic and social covariation is most adequately indicated by a construct which is not confined to a single linguistic rule or a sequence of rules but is instead a unique sociolinguistic construct. (1991:26)

Wolfram's comments are largely in response to the development of variable rules (Labov 1969), which tied sociolinguistic variables to individual rules in an SPE-style (Chomsky and Halle 1968) system of phonology. Thus a variable rule was a representation of the alternation within the grammar, and a locus for the application of non-categorical constraints (on which, see §4.2.4 below). Wolfram suggests that this conception of linguistic variables as a matter of (individual competence) grammar is a mistake.

However, before we throw the baby out with the bathwater, I would like to suggest that just because the alternation is not a matter of grammar doesn't mean that the constraints are also necessarily outside the grammar. The general picture I have in mind is this: The grammar provides choices in that it provides different ways of saying similar things. Such choices get associated with social differences or social meaning, and get exploited by speakers on the basis of that social meaning.[2] Non-categorical grammatical constraints are therefore associated directly with each variant, as factors that can affect how likely it is to be chosen.

To summarize, in most cases it is awkward to represent a linguistic variable in a grammar, but each variant is represented. Representing non-categorical constraints (discussed below) does not require a representation of the variable, as they can attach to each variant. Of course 'factors affecting how likely a variant is to be chosen' need not necessarily be a matter of grammar. The experiment reported in this chapter provides preliminary evidence that they are at least a matter of linguistic knowledge. The remainder of this section reviews the collection of variants that make up the variable of AAVE copula absence, and the constraints that affect their distribution.

4.2.2 The AAVE copula as a linguistic variable

The variable to be studied here is composed of a subset of instances of the copula in AAVE, in particular, those that appear in finite, non-past, non-habitual, non-first singular contexts. The forms that appear in those contexts are full forms (*is* and *are*), contracted forms (/z/ and /r/) and 'zero' forms, as illustrated in (1) to (2) above. In research on this variable, tokens that occur in past tense or non-finite contexts are excluded, as both contraction and the zero form are generally thought to be impossible in these environments.[3] Tokens that occur with first singular subjects are also excluded, as the zero form is again impossible (although both contracted and full forms are possible). Finally, 'habitual' or 'invariant' *be*, as in (4) from Wolfram (1969:180), is excluded on the grounds that it doesn't mean the same thing as the

[2] Such use constitutes a feedback loop wherein socially meaningful uses of variables are part of the construction of the meaning itself. See Eckert 2000, §5.1.2 of Chapter 5 and §6.2 of Chapter 6.

[3] See, however, the discussion of the zero copula and past tense in §3.3.2 of Chapter 3.

conjugated or zero forms.

(4) That's why I wonder why I don't see him—he usually be 'round.

Which of Wolfram's categories does this variable fall into? As a first pass, it appears to be an example of both 'structural category realization' and 'allo-forms within a structural category'. That is, the difference between the full and contracted forms would be a matter of allomorphy, while the zero form could be a matter of non-realization of the 'structural category' copula. However, there are still other ways of looking at it. Labov's (1969) analysis was in terms of two linguistic processes: deletion of the initial vowel of *is* or *are* (contraction) followed by deletion of the remaining consonant. The silent verb analysis in Chapter 3 suggests that the difference between the zero form and other forms of the copula is really a matter of lexical choice. The final constructional analysis of Chapter 3 makes it instead a matter of constructional choice (between an ordinary *head-comp-ph* and the *silent-verb-phrase*).

Labov's treatment of this variable as a phonological variable assumes semantic equivalence, as does a treatment in terms of straight allomorphy. However, if it is a matter of lexical or constructional choice, the possibility of semantic differences comes into play. In fact, there is some slight evidence for a semantic difference in terms of tense: Recall that, while the overt forms are marked for present tense, the zero form may in fact be unmarked for tense or underspecified between past and present tense, at least for some speakers. This is shown by sentences such as (5).[4]

(5) I dug being militant cause I was good. It was something I could do. Rap to people. Whip a righteous message on em. People knew my name. They'd listen. And I'd steady take care of business. **This was when Rap Brown and Stokeley and Bobby Seale and them ϕ on TV**. I identified with those cats...

There is also the possibility of semantic difference in the case of tokens with emphatic stress on the copula. Labov (1969) notes that this is one context with categorical copula presence (indeed, of full forms). As such, it should be excluded from studies of non-categorical constraints on variation. However, it is doubtful that

[4] From John Edgar Wideman's *Brothers and Keepers*, 1984, p. 114, cited in Green (in preparation).

all such tokens are successfully identified and eliminated in production studies of this variable. The perception experiment reported in this chapter contrasted only contracted and zero forms of the copula, largely because the full form introduces possible intonation differences as well as whatever semantic effect might follow from emphasis. The decision to contrast only zero and contracted forms in the experiment is supported by Wolfram's finding that the proportion of full forms was relatively constant across the four social groups of AAVE speakers that he studied in Detroit, ranging from 18.1% for the lower working class speakers to 27.4% for the upper middle class speakers. The white speakers from Detroit included in the study had 20.2% full forms. This contrasts with rates of zero forms ranging from 4.7% for the upper middle class speakers to 56.9% for the lower working class speakers. Wolfram concludes that "[t]he social significance of this variable is found in the different ratios of zero realization to contraction." (1969:168) In the rest of this chapter, I will refer to the zero variant as either the 'zero form' or as 'copula absence', using 'copula presence' or 'overt forms' to refer to either the contracted or full forms.

In order to consider copula presence/absence to be a variable of AAVE, it is necessary to establish that the forms with copula presence are not simply evidence of code-switching into a more standard variety. If they are, then the variation is not at the level of the form of the copula but rather at the level of choice of code. However, as discussed in Section 3.2 of Chapter 3, there is good reason to believe that both copula presence and copula absence are alternatives provided by the grammar of AAVE.

A final consideration with respect to this variable has to do with the conflation of copula and auxiliary uses of the verb *be*. That is, most studies of this variable do not exclude tokens on the basis of the syntactic category of the complement of *be*. Instances of the copula supporting nominal predicates are considered on a par with instances of the auxiliary before *ing*-form verbs and before *gon(na)*. As discussed in §4.2.4 below, the category of the complement is analyzed as a non-categorical constraint on the distribution of this variable, and it is precisely this constraint that I will be studying here. What motivates treating these different uses of *be* as one variable, when in other languages the copula is not also used as a tense auxiliary

(and in fact, different kinds of non-verbal predicates can require different copulas)? The main motivation is the forms that appear: when the copula is present (i.e., not zero), it takes the forms *is* and *are*, or /z/ and /r/, across all of the different following environments. (And the formal similarities of course hold in the other parts of the paradigm of *be*—past tense, *am*.) Further, AAVE and other varieties of English are not alone in using the same form across these categories. In their company, one finds Bengali and Nahuatl,[5] and if compound tenses other than progressives are allowed, several Romance languages as well. Finally, there is syntactic evidence (discussed in §2.5 of Chapter 2) for treating copula and auxiliary uses of *be* as the same lexeme.

4.2.3 Social constraints on the AAVE copula

For Weinreich et al., a linguistic variable is not merely a choice-point in a linguistic structure:

> A linguistic variable must be defined under strict conditions if it is to be a part of a linguistic structure; otherwise, one would simply be opening the door wide to rules in which "frequently," "occasionally," or "sometimes" apply. Quantitative evidence for *covariation* between the variable in question and some other linguistic or extralinguistic element provides a necessary condition for admitting such a structural unit. (1968:169)

This and the following subsection discuss some of the quantitative evidence pertaining to the covariation of forms of the copula in AAVE with other linguistic and extralinguistic elements.

As with most sociolinguistic variables, the realization of the AAVE copula is sensitive to some social variables and not to others. The most thorough investigation of the AAVE copula across different social groups is that of Wolfram 1969. The 48 speakers in this study were balanced across four class groups (lower working class, upper working class, lower middle class, and upper middle class), three age groups (10-12 years, 14-17 years, and 30-55 years), and both genders.[6] Wolfram found that the use

[5] John McWhorter, copula project files

[6] The interviews were carried out by researchers who were not African American. As Wolfram notes, this had the result of eliciting the style that these speakers use to address respected strangers.

of the zero forms correlated inversely with social class (working class speakers use more zero forms), gender (men use more zero forms) and age (children use more zero forms than teenagers or adults, with the exception of the lower working class group in which the teenagers had the same high rate of zero forms as the children). The one social variable in his study that did not correlate with copula absence/presence was racial isolation—the amount of contact a speaker had with white speakers.

Rickford and McNair-Knox (1994) consider another kind of social variable: addressee and topic-influenced style shift. They examined two interviews with the same speaker (Foxy Boston) in East Palo Alto, one by McNair-Knox who, in addition to being African American, was also a member of Foxy's community, and one by a European American interviewer. An additional difference was that McNair-Knox's daughter (who was a teenager, like Foxy) accompanied McNair-Knox on her interview. Foxy's use of several AAVE variables was sensitive to both the addressee and the topic of conversation. Her rate of copula absence was 70% with the African-American interviewer, compared with 40% with the European-American interviewer. Within the first interview, copula absence ranged from 43% in a discussion of graduation and college/career plans to 86% when the topic moved to boy-girl conflicts and relations. Within the second interview, the topic with the least copula absence was recreation at 9% while the topic with the most was "wives and slamming partners" (a discussion of boy-girl relationships) at 82%. College and career plans came in at 17% in this interview.

In summary, rates of copula absence in AAVE have been shown to vary with socioeconomic class, gender, age, topic of conversation, and the addressee.

4.2.4 Grammatical constraints on the AAVE copula

The realization of the AAVE copula is also sensitive to a number of grammatical factors, and some of these effects have been remarkably consistent across studies in different communities. Labov's (1969) study of copula absence among young speakers in South Harlem identified effects of the type of subject (pronoun favors copula absence, full NP favors copula presence), the following grammatical environment (i.e., the part of speech of the predicate, discussed in detail below), and the preceding phonological environment (a preceding vowel favors contraction and a preceding consonant favors the zero form). These constraints differ from those discussed in Chapter 3 in that they are not absolute: copula absence is not ungrammatical but only uncommon in a disfavoring grammatical environment. Therefore I will be referring to these constraints as non-categorical (grammatical) constraints on AAVE copula absence.

The effect of a preceding NP vs. pronoun was confirmed in Detroit by Wolfram (1969), in Los Angeles by Baugh (1979) and in East Palo Alto by Rickford et al. (1991). The effect of the preceding phonological environment is less straightforward. As discussed in Section 3.3.3 of Chapter 3, Rickford et al. (1991) show that it is dependent on how one calculates the rates of each variant.

Wolfram (1969) found a further effect of person and number, comparing *is* deletion to *are* deletion.[7] (Labov (1969) only studied *is* deletion.) *Are* environments favor the zero form more strongly than *is* environments do in Wolfram's data. Bailey and Maynor (1987) confirm this pattern with data from children and 'folk speakers' (adults over 70 with less than a grade school education) in East-Central Texas. However, in Rickford et al.'s (1991) East Palo Alto data, the effect is reversed, with *is* environments favoring the zero form and *are* environments favoring overt forms. This difference could have to do with r-lessness. Wolfram's speakers have variable r-lessness, and the speakers with more r-lessness (working class) also have higher rates of *are* absence and bigger absolute differences in their rates of *are* absence and *is* absence. Bailey and Maynor (1987) and Rickford et al. (1991) do not indicate whether their speakers have r-less or r-ful dialects, but it is likely that Bailey and Maynor's speakers had r-lessness while Rickford et al.'s did not.

With so many factors, it is difficult to be certain of the exact contribution of each one by eye-balling frequency data. The form of statistical analysis called Varbrul (which runs a step-wise multiple regression) was developed to answer this kind of problem (by Cedergren (1973), Sankoff (1975), and Rousseau and Sankoff (1978)).

[7] The scare quotes are there because AAVE does allow *is* with non-third singular subjects, so that when plural subjects co-occur with the zero form, it's not clear that *are* would be the overt form. However, Wolfram reports that in his data, such non-agreement cases account for less than 5% of the overt forms.

Varbrul ascertains the effects of the various independent variables on the distribution of the dependent variable. In particular, it returns an input value (the general propensity that the speakers had for producing the distinguished variant) and factor weight for each factor in each factor group that was found to have a statistically significant effect. For example, the type of subject forms a factor group in studies of AAVE copula absence. The factors are NP subject and personal pronoun subject. A factor weight of above .5 (e.g., the factor weight of .62 that Rickford et al. (1991) found for the factor personal pronoun[8]) indicates that the factor favors the distinguished variant (here, copula absence). Similarly, a factor weight of below .5 (e.g., Rickford et al.'s (1991) factor weight of .42 for NP subject) indicates that the factor disfavors the distinguished variant. The higher/lower the factor weight, the stronger the favoring/disfavoring effect.[9]

The single most widely studied non-categorical grammatical constraint on copula absence, and the one that is the focus of the experiment, is the effect of the following grammatical environment. Table 4.1, excerpted from Rickford's Table 6.16 (1998:190), provides a summary of nine data sets from studies of copula absence in various AAVE speaking communities. The left half of the table gives information about the data set: whether it concerns *is*, *are*, or both; the population studied; and the author. The right hand side of the table gives information about the effect of following grammatical environment on copula absence. With the exception of Wolfram's study (#4), the values are Varbrul weights, with copula absence as the distinguished variant. Wolfram's results were reported are raw frequency data, given in percentages. The five environments considered are noun phrase (__NP), locative (__Loc), adjective (__Adj), *-in(g)* form verb (__V+ing), and *gon(na)* (__gon). These environments are arranged according to the ranking in (6), where environments to the left are the least favorable to copula absences and the environments to the right are the most favorable. (7) to (11) illustrate each of the environments with examples from Wolfram 1969.

(6) __NP < __Loc < __Adj < __V+ing < __gon

[8] Table 6, p. 117, Straight Deletion, *is* and *are* combined, input probability .35.
[9] For more on Varbrul, see Guy 1988.

Studies				Environments				
#	Form	Place	Source	__NP	__Loc	__Adj	__V+ing	__gon
1	*is*	NYC Thunderbirds	Labov 1969	.2	.36	.48	.66	.88
2	*is*	NYC Jets	Labov 1969	.32	.52	.36	.74	.93
3	*is*	NYC Cobras	Baugh 1979	.14	.31	.72	.59	.78
4	*is+are*	Detroit WC	Wolfram 1969	37%	44%	47%	50%	79%
5	*is*	LA	Baugh 1979	.32	.29	.56	.66	.69
6	*are*	LA	Baugh 1979	.25	.69	.35	.62	.64
7	*is+are*	Texas kids	Bailey & Maynor 1987	.12	.19	.25	.41	.89
8	*is+are*	Texas adults	Bailey & Maynor 1987	.09	.15	.14	.73	.68
9	*is+are*	EPA	Rickford et al. 1991	.29	.42	.47	.66	.77

Table 4.1: Copula absence in AAVE in different communities, by following grammatical environment. (Values given are Varbrul factor weights, with the exception of the data from Wolfram (1969), which is given in percentages.)

(7) __NP: She a nurse.

(8) __Loc: They out there in space.

(9) __Adj: She real nice.

(10) __V+ing: Do anything if you fighting.

(11) __gon: I really don't think John gonna make it.

While the pattern of non-categorical constraints in (6) is robust, it is not perfectly consistent across all of the studies. Four of the nine data sets (#s 1,4,7,9) do have exactly this pattern. Of the others, three are off by the ordering of one pair of environments: in #2, __Loc shows more copula absence than __Adj, in #3, __Adj shows more copula absence than __V+ing, and in #5, __NP shows more copula absence than __Loc. The remaining two data sets have more extensive differences. In #6, __Loc has the most copula absence, skipping ahead three slots in the ranking as compared to the most common pattern. In #8, __Loc shows more copula absence

Pair	n	Pair	n
NP < Loc	8	Loc < V+ing	8
NP < Adj	9	Loc < gon	8
NP < V+ing	9	Adj < V+ing	8
NP < gon	9	Adj < gon	9
Loc < Adj	6	V+ing < gon	8

Table 4.2: Pairwise rankings of environments

than __Adj (although they are extremely close), and __V+ing shows more copula absence than __gon.

It is important to note that these data are not exactly comparable. Some of the studies considered *is* only or *is* and *are* separately, while others treated *is* and *are* together. It is perhaps noteworthy that three of the four data sets exhibiting the ranking in (6) are of the *is+are* type, and three of the four data sets of the *is+are* type show the same pattern.[10]

Also, while the nine data sets do not all have the same ranking, they are still similar. One way to look at the similarity of these patterns is by pairwise ranking. Table 4.2 gives the number of data sets that conform to each of the pairwise rankings in the ranking given in (6). Only the ranking Loc < Adj has as few as 6 of the 9 datasets attesting it. All of the others have either 8 or 9. Looked at this way, the pattern appears quite robust.

Within each of these studies, the pattern found was argued to be robust (statistically significant)—i.e., something more than an accident of which data happened to be recorded. That (roughly) the same pattern is found across many communities is a strong confirmation. This pattern of non-categorical constraints on copula absence appears to be a property of AAVE, one that must either follow from something else in the language, follow from functional constraints on performance, or be learned, for there is no other way for it to be reproduced in community after community.

There is one important caveat to this argument: just because an effect can be

[10]Rickford et al. (1991) compare the results *is* and *are* deletion in their study and conclude that the constraints are similar enough in both cases to warrant treating them together as one variable.

found by considering the following grammatical environment doesn't mean that that's really what's going on. It could be that the correlation is between semantic properties and rates of copula absence, and that the strength of this effect together with a reasonably close pairing of syntax and semantics are enough to let the effect shine through even though we've been counting the 'wrong' thing. One semantic distinction that may be having an effect is that between individual and stage level predicates. Individual level predicates (e.g., *tall* in *Kim tall*) denote properties that do not change easily over the course of an entity's existence. Stage level predicates (e.g., *busy* in *Kim busy*) denote properties that tend to be temporary, i.e., are true of 'stages' of an individual. Another semantic distinction to investigate is the subdivision of NP predicates into identity sentences (e.g., *Samuel Clemens Mark Twain*) and predicational sentences (e.g., *Kim a doctor*). Both of these distinctions would seem to be equally likely candidates for non-categorical constraints as the syntactic distinctions studied so far. However, semantic constraints would still be grammatical constraints. Further, it seems unlikely that a reanalysis of the production data in terms of semantic categories would turn up only categorical constraints, given minimal sets such as in (1–2). Although the syntactic effect shown in Table 4.1 may only be a reflection of a semantic effect, the fact remains that it is robust. Syntactic or semantic, it must therefore either be a part of the grammar of AAVE or follow from something else in the grammar or in performance.

4.3 Hypothesis

This leaves us with the question of what (if anything) in the grammar of AAVE underlies the synchronic pattern of copula absence. The various existing theories of non-categorical constraints on variation will be discussed at length in Chapter 5. In brief, there are three possibilities in the literature: Variable rule accounts (e.g., Labov 1969) posit optional context-sensitive rules as part of the grammar. These optional rules have a probability of application that is affected by the grammatical context. Optimality Theoretic (OT) accounts (e.g., Anttila 1997) posit partially unranked grammars. The distribution of a variant (or candidate, in OT terms) is determined by the number of complete rankings that select that variant. The third possibility is functional accounts (e.g., Kiparsky 1972) where the disfavoring effect of certain grammatical environments is attributed to functional pressures on performance. The variable rules model, but not the others, requires that speakers have direct (if tacit) knowledge of non-categorical constraints on variation.

If speakers have knowledge of non-categorical constraints on variation, it is almost certainly tacit and inaccessible to introspection. Fortunately, Labov's (1963) finding that sociolinguistic variation is socially meaningful provides a jumping off point for constructing an experiment. If sociolinguistic variation is socially meaningful, then the social value of variants might interact with the non-categorical constraints. In particular, Wolfram makes the following observation concerning the data in Table 4.3 (from Wolfram 1969:172):[11]

> The relatively high frequency with which zero realization is found preceding intentional future *gonna* among middle-class informants suggests that zero realization preceding *gonna* is less stigmatized than zero realization in other environments. (1969:172–173)

	Following grammatical environment				
Speaker group	__Adj	__NP	__Loc	__V+ing	__gon
Middle class	1.6%	4.2%	13.3%	11.3%	33.3%
Working class	36.5%	47.3%	44.4%	50.0%	78.9%

Table 4.3: Wolfram's copula absence by following grammatical environment and social class

Generalizing to allow for social values other than stigmatization, I propose the following two-part hypothesis:

I Copula absence/presence in AAVE is associated with some social value.

[11]See also Sylva and Zwicky's (1975) observation that the stylistically marked syntactic rules they discuss can be more or less stylistically marked depending on the lexical items they interact with. For example, they observe that existential *there* is relatively 'formal', but only with verbs other than *be*.

II Copula absence/presence in AAVE is more strongly associated with that social value the more marked the environment is for each variant.

For example, if copula absence sounds confident, then copula absence before a noun should sound especially confident and copula absence before a verb somewhat less so.

Note that Part II of this hypothesis entails that speakers have knowledge of non-categorical constraints. However, what is at issue here is more than just the knowledge of the constraint. Previous formal approaches to constraints on variation, if they consider social constraints at all, treat them as separate from grammatical constraints. The hypothesis being tested here is that they interact. That is, social constraints are conceptualized as the social meaning of the variable, and grammatical constraints as the intensifying or attenuating effect of the grammatical environment on the social meaning or social value of the variable.

4.4 Methodology

The experimental design was based on the matched-guise methodology of Lambert et al. 1975.[12] Lambert et al. were looking at language attitudes in a French-American community, and in particular the social evaluation of speakers based on language variety. The varieties represented in the study were European French, middle class Canadian French, lower class Canadian French, middle class Madawaskan (local) French, lower class Madawaskan French, middle class Madawaskan English and middle class non-regional English.

They presented to listeners from various groups within the community recordings of the same passage spoken in the different language varieties. In some cases, bilingual or bidialectal talkers[13] rendered multiple versions of the passages, i.e., appeared in

[12]See also Lambert 1967, Giles 1971, Anisfeld and Lambert 1964, Lambert et al. 1966, S. Lambert 1973, and Lambert and Tucker 1975.

[13]I will use the term 'talker' for the speakers who made the recordings in a matched-guise experiment because the term 'speaker' is used to refer to a person with competence in a given language, regardless of what s/he is using the competence to do in the situation being considered. The participants who listened to the recorded stimuli will be referred to as 'listeners' in general, but also as 'speakers' when it is important to highlight their competence of a particular variety.

multiple guises in the study.[14] The talkers each read the same 20-second passage from *Le Petit Prince* by Saint-Exupéry, translated or adapted into the different varieties. Lambert et al. made their experimental tape by repeating each speech sample end-to-end, and randomly ordering the resulting 40-second speech samples.

They then presented this recording to listeners, and asked them to rate talkers on 20 7-point scales, to rate (again on 7-point scales) how much they felt they resembled the talkers and how much they would like to be like the talkers, and to guess at the occupation of the talker. Table 4.4 lists the scales used by Lambert et al. in their studies with high-school and college-age participants. These scales were chosen on the basis of preliminary studies with the target populations aimed at discovering which personality qualities were valued.

1.	intelligent – not-intelligent	11.	unsociable – sociable
2.	active – passive	12.	short – tall
3.	unfair – fair	13.	ambitious – not ambitious
4.	truthful – untruthful	14.	friendly – unfriendly
5.	good-looking – ugly	15.	not religious – religious
6.	not comical – comical	16.	strong – weak
7.	not courageous – courageous	17.	impolite – polite
8.	unsure – confident	18.	happy – sad
9.	likeable – hateful	19.	selfish – not selfish
10.	reliable – unreliable	20.	determined – not determined

Table 4.4: Rating scales used by Lambert et al.

Lambert et al. found that their listeners have definite social evaluations of the different language varieties. For example, the high school students in their study middle class Madawaskan French as favorably as European French and the English varieties on the traits that Lambert et al. related to social attractiveness (fair, likeable), but not as favorably on scales related to competence (intelligent, determined).

For the experiment reported in this chapter, I extended the matched-guise methodology to test the social evaluation of an individual linguistic feature. This entailed

[14]In some of the earlier studies (Lambert 1967, Giles 1971) the methodology was purely matched-guise, in the sense that each talker recorded the passage in each variety being considered.

certain modifications: a drastic reduction in the length of the each speech sample, and a corresponding reduction in the number of scales used. These changes are described in further detail and motivated in the following subsection.

4.4.1 Stimuli

Test sentences

As the purpose of this study was to test the social evaluation of a particular linguistic feature (copula absence) and the effect of the grammatical context on that evaluation, the speech samples had to be much shorter than those used in Lambert et al.'s study. This is because the longer the speech sample, the more likely the introduction of confounding variables. Similarly, the sentences were designed to avoid AAVE-and SAE-specific features as much as possible. Even if held constant across all of these stimuli, a feature other than copula presence/absence that is strongly stereotyped could have flooded the effect of the variable under consideration. Nonetheless, since I was looking at the evaluation of copula absence within the system of AAVE, it was important to choose sentences that are a part of AAVE, and to make sure that their actual production conformed to AAVE phonology, etc. Further, in testing for the effect of the non-categorical constraint in question, it was important to control for all other known non-categorical constraints.

The non-categorical constraint to be tested is the effect of the following grammatical environment. Recall that production studies have established that the ordering of following grammatical environments, from least favoring of copula absence to most favoring, is as in (12).

(12) __NP < __Loc < __Adj < __V+ing < __gon

Ideally, one would want to test all of the environments to see if they are ordered in perception as they are in production. However, in order to keep the experimental task reasonably short, only one pair was tested. *Gon* is somewhat problematic because it is itself an AAVE-specific feature and the high rates of copula absence before *gon* suggest that it may be in the process of being reanalyzed as a modal.[15] Leaving out *gon*, the V+ing and NP environments allow for the greatest contrast.

The test sentences, designed to meet all of these considerations, are as follows:[16]

(13) a. Yeah I know her. She's teachin me piano at Music World.
b. Yeah I know her. She's my piano teacher at Music World.
c. Yeah I know her. She teachin me piano at Music World.
d. Yeah I know her. She my piano teacher at Music World.

Sentences (13a) and (13b) represent the copula presence condition, while (13c) and (13d) represent the copula absence condition. (13a) and (13c) represent the V+ing condition, and (13b) and (13d) represent the NP condition.

The sentences are about as close as they can be in phonological and semantic content given the requirement of setting up the different conditions, and are constant with respect to all other known constraints: the subject is always a pronoun, and always phonologically identical. There is no known effect of following stop vs. nasal.

The initial and final phrases (*Yeah I know her* and *at Music World*) are included to give the sentences some length.

In addition to the test sentences, the following sentences were also recorded as filler stimuli:

(14) a. Yeah I know her. She useta teach me piano at Music World.
b. Yeah I know her. She useta be my piano teacher at Music World.
c. Yeah I know her. She taught me piano at Music World.
d. Yeah I know her. She was my piano teacher at Music World.

The filler sentences match the test sentences fairly closely in their semantics, but not as closely as the test sentences match each other. Likewise, there is more variation in phonological content.

The talkers that I recruited to record the stimuli confirmed that all of these sentences would sound natural in the middle of an AAVE conversation.

[15] Although not one with all of the properties of a true auxiliary: for example, it doesn't invert. (Salikoko Mufwene, p.c., Jan., 2000). Perhaps it is more like *better* in SAE *You better not do that.*

[16] Ideally, such an experiment would involve multiple similar test sets. However, in order to keep the listeners' task to a manageable length, only the test set in (13) was used.

Recording the stimuli

Eight talkers were recruited from the Stanford community to record the sentences. All were (at least) bidialectal, commanding both AAVE and SAE.[17] All were female. While it would certainly be interesting to look into the effect of gender of the talker on the evaluation of copula absence, gender was excluded as a variable in this study. The primary reason for this is that including both male and female talkers in the stimulus set would have made it easier for listeners to recognize individual talkers when they heard them for a second, third or fourth time. This could have led to listeners trying to be consistent in the way they rated different talkers.

The sentences were presented to the talkers in written form. This immediately raised the issue of how to avoid reading style and get something that more closely approximated natural AAVE speech. The issue was further exacerbated by the presence of the white researcher, the use of recording equipment, and the surroundings: the sound-proof room in the Stanford phonetics laboratory.

The solution adopted was to have the talkers do the recording session together in groups as much as possible. In the initial session, three talkers were brought together. They first read the sentences out loud to each other and discussed different ways of saying them. I tried to background myself as much as possible during this interaction. Then the talkers took turns recording all eight sentences (i.e., both the test and filler sentences) sometimes just saying them all in a row, sometimes with a prompt from another one of the talkers ("You know her?"). Subsequent talkers came in pairs or alone, and were all played the existing tape to give them an idea of how the initial group had decided the sentences should sound. This method succeeded in producing relatively natural sounding recordings despite the conditions and in minimizing the variation across talkers in features such as intonation. It is important to emphasize at this point that both sentences with copula absence and sentences with copula presence are possible AAVE sentences, and were produced with similar intonation and other phonological features. That is, it is NOT the case that the copula presence sentences represent standard English while the copula absence sentences represent AAVE.

[17] I did not systematically collect geographical data on the talkers. Although all were currently living in California, some at least were from other regions of the country.

All of the talkers were recorded saying each sentence multiple times. Each instance was digitized using SoundEdit on a Macintosh and the actual stimuli were selected on the basis of clarity, lack of hesitation or extraneous sounds like laughter, and consistency of intonation across the different sentences.[18] The four talkers with the best recordings were then designated as test-talkers and the other four as filler-talkers. The test-sentence recordings from the test-talkers and the filler-sentence recordings from the filler-talkers were combined to make 32 stimuli.

Finally, in order to avoid any effect of extraneous variation, a recording of *Yeah I know her* was selected for each talker from one of the recordings not chosen as stimuli and pasted onto the beginning of each of that talker's test or filler stimuli followed by a standardized pause of 0.15 seconds. The *Yeah I know her* was chosen to prosodically match the the rest of the stimuli as closely as possible and so usually came from the same reading of the sentences as the stimuli (e.g., from a filler sentence for a test-talker and vice versa). This added a certain unnaturalness to the stimuli, but it was equally unnatural for each stimulus. This was possible because the *Yeah I know her* part was prosodically independent of the rest of the sentence in all of the recordings chosen as stimuli. To do the same thing with the end of the sentences (*at Music World*) would have introduced too much unnaturalness.

The resulting test stimuli ranged from 2.482 seconds to 2.698 seconds in length, with an average of 2.592 seconds. These lengths include 0.1 seconds of silence at the beginning and end of each stimulus. The filler stimuli ranged from 2.36 to 3.208 seconds, with an average of 2.711 seconds. It was decided not to follow Lambert et al.'s strategy of lengthening the stimuli by repeating them on the grounds that this would give listeners more of a chance to consciously identify the variation in the realization of the copula.

Test scales

Lambert et al. were researching the language attitudes of members of the community they studied, and indirectly the ethnic identification of people growing up in a bicultural environment. Accordingly, they used the scales to assess the stigma or prestige

[18] All judged impressionistically by the experimenter.

associated with each language variety represented in their recordings. In their analysis, they group the scales into subgroups associated with competence (intelligence, determination, etc.) and social attractiveness (fair, likeable, etc.).

Since I was focusing on one feature (copula absence) there was the possibility that it has some more specific social value, akin to the way *ing*/*in* is associated with formality/informality (Fischer 1958), and that one of the scales chosen would be close enough to this social value to reflect it. This argues for including as many scales as possible. On the other hand, too many scales would end up making the experiment too long and could also lead to the last several scales being presented too long after the listener heard the stimulus. In the end, I settled on the seven scales listed in Table 4.5.

1.	comical – not comical
2.	confident – not confident
3.	well educated – not well educated
4.	good job – not a good job
5.	likeable – not likeable
6.	polite – impolite
7.	reliable – unreliable

Table 4.5: Scales used for the experiment

Each scale was presented with a question for context: "How likeable/reliable/well educated does this person sound?" "How polite/comical is this person being?" "How confident do you think this person is feeling?" and "How good of a job do you think this person has?"[19] All of the scales had seven points from which the listeners could choose.

These particular scales were chosen so as to cover as much ground as possible in the hopes that at least one of them would be relevant to copula absence. Further, they were meant to be traits that are not necessarily correlated. In particular, someone could sound confident without sounding well educated. Also, the scale 'reliable –

[19] The form of this last question is not grammatical for all speakers. Whether it was grammatical for all of the speakers in this study and, if not, whether it would have had any effect on the results, is unclear.

unreliable' was included in the hopes that it would index in-groupness, that the listeners would interpret it as reliable in the sense of 'a friend you can count on'.

Language attitudes questionnaire

Lambert et al. used the ratings as a measure of their participants' attitudes towards the language varieties they were considering. In this study, I was investigating the social evaluation of a particular feature and the effect of the surrounding grammatical environment on that feature. Overall attitudes towards AAVE certainly have an effect on the evaluation of AAVE features. Combining ratings from speakers with a positive stance towards AAVE and from speakers with a negative stance could potentially mask interesting effects. Therefore, a language attitudes questionnaire was included in the study.

The questionnaire consisted of five statements. Listeners were asked to specify whether they agreed or disagreed with each statement, and to what extent. The five statements are given in Table 4.6.[20] The word AAVE was explained in the instructions for this questionnaire.

S1	There are some situations in which it is better to speak AAVE than Standard English.
S2	A Black person who doesn't speak AAVE isn't really Black.
S3	AAVE is bad English and should never be used in any situation.
S4	It would be weird for me to speak Standard English instead of AAVE with my Black friends.
S5	When I hear a stranger speaking AAVE, I assume they are less educated.

Table 4.6: Statements for language attitudes questionnaire

[20]These statements were designed for an intended participant pool of African Americans, and were somewhat awkward for the non-African-American participants.

4.4.2 Participants

Participants were recruited from a introductory psychology course at a community college in a community in California with a large African-American population. Participants earned extra credit in their course for their participation and were also paid $5.

Since participants could earn extra credit, the experiment had to be open to anyone in the class. The people who chose to participate fell into the five ethnically and linguistically defined groups given in Table 4.7. These groups will form the basis of the analysis below, and will be referred to by the numbers assigned to them in Table 4.7.

I	AAVE speakers (all African American)	N = 11
II	African Americans who do not identify as AAVE speakers, but who are familiar with AAVE	N = 5
III	participants who are familiar with AAVE but are not African American	N = 6
IV	native speakers of English with no familiarity with AAVE	N = 6
V	non-native speakers of English with no familiarity with AAVE	N = 7

Table 4.7: Groups of listeners

Familiarity with AAVE was determined in an exit interview as described in §4.4.3 below. Whether or not a person spoke (any variety of) English natively was determined on the basis of the demographic questionnaire. Any listener who grew up in a non-English-speaking country was counted as non-native. Note that membership in Groups II and III only requires familiarity with AAVE and not native-speaker status in English. In fact, one listener in Group III grew up in the Philippines, and a speaker who grew up in Haiti was included in Group II.[21] Groups II and III are distinguished on the basis that the African Americans' experience with AAVE is substantially different from the others'. Group III listeners reported that they were familiar with AAVE because they went to predominantly African American high schools and/or participate in Hip Hop culture.

[21]This speaker learned both Standard English and AAVE in this country. He was included in Group II on the basis that while he didn't have native command of AAVE, he did have extensive experience with it. His responses also pattern with those of the other speakers in Group II.

The demographic questionnaire consisted of the questions listed in Table 4.8. The ethnicity question was added only in the second session, as I had originally expected only to recruit African American participants. Participants' first answers were accepted, so that if father's occupation was reported as "retired" or "deceased" no further information was collected unless volunteered by the participant.

Are you (male/female)?
What is your current occupation?
What was your previous occupation (if any)?
What is your father's occupation?
What is your mother's occupation?
Where did you live between the ages of 2 and 18?
When did you move to the Bay Area?
How old are you?
Ethnicity

Table 4.8: Demographic questionnaire

The demographic information is summarized in the next few tables. Table 4.9 summarizes the information on age and gender by group. Table 4.10 gives the ethnic self-identification of the participants as well as information on where they grew up.[22]

Group	age range	median age	# men	# women
I	19–55	26	3	8
II	24–33	26	2	3
III	18–27	20.5	1	5
IV	19–40	25	2	4
V	18–34	21	2	5

Table 4.9: Age and gender information, by group

[22] Some of the African American participants gave their ethnicity as 'Black' and some as 'African American'. However, the software used for this portion of the experiment allowed participants to see the responses of the previous participant, and they often seemed to choose based on what was already there. That is, if the previous participant had said African American, another African American participant would say the same. If the previous participant was, say, Mexican American, an African American participant might say either 'Black' or 'African American'.

Group	ethnic self-identification	home state/country[a]	N
I	(didn't ask)	California	4
	African American	California	3
	Black	California	2
	African American	Florida	1
	Black	Louisiana, California	1
II	African American	California	2
	African American	Alabama	1
	Mixed	California	1
	Black	Haiti	1
III	White	California	1
	Jewish	Texas	1
	Filipino	Philippines, California	1
	White	California	1
	Caucasian	Ohio, New Mexico, California	1
	Asian	Illinois, California	1
IV	(didn't ask)	California	2
	Indian	US Virgin Islands, British Virgin Islands, California	1
	Latina	California	1
	White	California	1
	Mexican American	California	1
V	Chinese	China, California	3
	Asian Pacific	Korea	1
	Asian	China	1
	Vietnamese	Vietnam, California	1
	African American	Nigeria	1

[a] Where the listener lived between the ages of 2 and 18

Table 4.10: Ethnicity and home state/country, by group

Tables 4.11 and 4.12 give the responses to the questions about occupation, which were included to gather information about the socio-economic status of the participants. Rather than summarize this information based on a scale such as that used by Labov (1966), I have given the actual answers from each participant. The main reason for this is that these data were not collected in a consistent manner: in many cases participants gave job titles from which it is difficult to deduce occupations, or simply employers ("Pac Bell") or industries ("fast food"). There are further reasons that these responses are probably not reliable indicators of socio-economic status: Some of the participants' parents probably still live in other countries (notably, participant N in Group II and some of the participants in Group V), where the status of different occupations may not be the same as in the US. A second consideration is the age range of the participants. For the younger participants, occupations such as food service are more likely to be temporary part-time jobs. Finally, there are many blank cells in this table. In many cases, the occupation of a parent was unknown, or simply given as "retired" or "deceased". Blank cells in the second column (previous occupation) indicate that the current occupation is the participant's first.

L	Current Occupation	Previous Occupation	Father's Occupation	Mother's Occupation
Group I				
A	Bank teller	Student	Clerk	Nurse
B	Receptionist	Clerk		Security
C	Resident Management Officer	Program Monitor	Welder	Licensed Vocational nurse
D	Hair stylist			Pac Bell
E	Mortgage Consultant	Student	Coach	Psychologist
F	File clerk	Material handler		Computer programmer
G	Dispatcher	Certified nurse's assistant		Nurse
H	Executive asst.	Principle Secretary	Deputy sheriff	Job developer
I	Medical asst.			
J	Student	Accounting	Police man	Housewife
K	House keeping	Dishwasher		House keeping
Group II				
L	Auditor	Teller		
M	Industrial sales	Industrial management	Industrial work	Social work
N	Security	Student	Farmer	Housewife
O	Meeting planner	Administrative assistant	(pilot?)	Administrative assistant
P	State employee			County employee

Table 4.11: Current, former, mother's and father's occupation, Groups I-II

L	Current Occupation	Previous Occupation	Father's Occupation	Mother's Occupation
Group III				
Q	Student	Pharmacy intern	Janitor	
R	Student		Manager	School teacher
S	Student	Waitress	Lawyer	Housewife
T	Student	Phone operator	City manager	Paralegal
U	Student		Engineer	Housewife
V	Medical coordinator	Cashier	Teacher	Housewife
Group IV				
W	Human resources coordinator	Insurance representative		Instructional asst. (elementary school)
X	Sales		Doctor	Doctor's asst.
Y	Student		Engineer	Cosmetology
Z	Student		Teacher	Accountant
a	Student		House painter	USPS
b	Stay at home mom	Appointment setter	Chemical engineer	Works for social secretary
Group V				
c	Nurse		Doctor	Nurse
d	Student		Food service worker	Housewife
e	Student		Construction worker	Homemaker
f	Student	Fast food	Engineer (car design)	Home support service
g	Food service Student	Environmental services (water sampling)	Food service	Clothes manufacturing
h	Sales associate	Waitress	Cook	Cook
i	Student		Manager	works at a jewelry store

Table 4.12: Current, former, mother's and father's occupation, Groups III-V

4.4.3 Running of the experiment

Lambert et al. give the following description of how they introduced the task:

> [L]isteners were led to believe that they were being tested on their ability to infer the personality characteristics of a speaker using speech cues only; the language spoken or the dialect variation heard was said to be insignificant and to be overlooked; they were told that the task was analogous to estimating another's personality from a phone conversation or from overhearing but not seeing a communicator. (1975:135)

This introduction of the task requires participants to believe that there is some way that the experimenters were able to objectively determine the personality characteristics of the talkers so as to have something to compare the listeners' responses to. In my estimation, it is quite likely that the community college students I was recruiting as participants as well as the college students in Lambert et al.'s study would be suspicious of this.

Instead, I introduced the study by saying that I was researching how consistently people judge other people on the basis of their voice. I went on to say that I was particularly interested in how this worked in the AAVE-speaking community. I described AAVE as the language variety used by some African Americans in some particularly informal situations, and emphasized that not all African Americans speak AAVE, nor do AAVE speakers always use AAVE, most also speaking at least one other variety. I particularly encouraged anyone who spoke AAVE to participate, but welcomed anyone in the class as non-AAVE speakers who would form my control group(s). This introduction was given when I went to the psychology class, and given the patterns of attendance, it is not certain that all of the people who eventually participated were there to hear it.

The experiment took place in an instructor's office at the community college, with one to two participants at a time (there were two computers). The session began with the participants reading and signing a consent form. For any participants who missed the in-class introduction, the consent form provided the first information on the objective of the study:

> You are invited to participate in a research study on how consistently people judge other people on the basis of their voices. First you will be asked some questions about yourself (such as your age and occupation). Then you will be asked to listen to 32 short recordings of people speaking and rate each speaker on 7 scales representing personality characteristics, such as reliable/unreliable or polite/impolite. When you are done, you will be asked to complete a survey about different speech styles.

Once the demographic questionnaire was completed, the participants donned headphones attached to the computers they would use. The stimuli were presented to the subjects by the program PsyScope (Cohen et al. 1993), beginning with the instructions shown in Figure 4.1.

In this experiment, you will hear 32 sentence spoken by different people. After each sentence, you will be presented with 7 questions on the screen. Each question will have a scale to use in your answer. For instance, if the question is "How kind does this person sound?", the scale might look like this:

Kind • • • • • • • Unkind

If you think the person sounded very kind, click on the point closest to kind. If you think they sounded neither particularly kind nor unkind, click on the point in the middle, etc.

As soon as you click on a point, the scale will go away and you will be presented with the next scale. When you have finished all 7 scales, you will hear the next voice.

If you don't want to answer one of the questions, you can wait for it to time out.

The first trial will be a practice trial. When you are ready to start the practice trial, click the mouse button.

Figure 4.1: Instructions for main experimental task

The participants completed one practice trial and then were given a chance to ask questions before beginning the actual trials.

Each trial began with the message "loading speaker" being displayed for 0.5 seconds to give the participants a signal that they should get ready to listen. Then the screen went blank and one of the 32 test or filler stimuli was played over the headphones. The stimuli were presented in a near-random order: it was important to minimize the probability that stimuli from the same talker would be presented one after the other. The actual orderings were produced by randomly selecting one stimulus from each talker and presenting those eight stimuli in a random order before selecting another stimulus from each talker, etc. While this did lead to back-to-back stimuli from the same talker on occasion, it was less frequent than it would have been with a true random order. The order of presentation was determined separately for each listener.

Each recording was followed by the seven scales. The scales were presented in a random order, with the position of the positive and negative poles selected randomly (i.e., sometimes the negative adjective was on the left, and sometimes it was on the right). The presentation of the scales, together with the contextualizing question, was as in Figure 4.2.

How good of a job do you think this person has?

Good job • • • • • • • Not a good job

Figure 4.2: Presentation of scales

The question and scale remained on the screen until the participant put the mouse on or near one of the seven bullets and clicked,[23] or until the question timed out at 20 seconds. The timeout was included as a mechanism for allowing participants to

[23] Participants were able to use either a track pad or an external mouse.

pass on individual questions, as specified in the consent form. A click on or near the bullet closest to the negative end of the scale (e.g., 'not a good job') was coded automatically by PsyScope as 1, the next one over as 2, etc. After all seven scales were presented in this fashion, the next trial began with the "Loading speaker..." message.

The last part of the experiment to be presented on the computer was the language attitudes questionnaire. It was prefaced by another instructions screen, shown in Figure 4.3. Each of the five questions in Table 4.6 above was presented followed by the scale of responses shown in Figure 4.3. As with the rating scales, each statement and response scale stayed on the screen until either the participant clicked on one of the responses or it timed out. The time out period for these questions was originally set at 20 seconds, but extended to 50 seconds after some of the listeners who participated on the first day found 20 seconds to be too short.

Finally, when the computer-based portion was complete, there was an exit interview in which I informed the participants of the purpose of the study and determined their level of familiarity with AAVE. Whenever possible, the exit interview was done with two participants at a time. The participants were shown a list of the eight sentences (test and filler) and the exit interview proceeded roughly according to the following script:

> These are the sentences that the people were saying. How many different speakers did you think there were? ... Actually, there were eight speakers. Four said each of these sentences [the test sentences] and four said each of these sentences [the filler sentences]. These sentences [the test sentences] are the ones I'm interested in. The rest are filler sentences. As you can see, these two have the verb *is* in them and these two don't. In AAVE, both ways of saying it are grammatical. I was interested to see if the choice between saying *is* and leaving it out would affect how people rated the speakers.
>
> Also, in these two sentences the predicate is a verb, *teachin me piano*, while in these two it's a noun, *my piano teacher*. When you record people speaking AAVE, it turns out that they're much more likely to leave the *is* out when the predicate is a verb, like *teachin*, than when it is a noun, like *piano teacher*. So I am interested to see if that will have any effect on how people rate the speakers. My hypothesis was that if, for example, it

---

The second part of this experiment is a quick survey about your opinions about the language variety spoken by many African Americans in informal situations. Linguists call this variety African American Vernacular English, or AAVE for short. Other names include Black English, Ebonics, and Jive.

In this survey, you will be presented with 5 statements. Each statement will be accompanied by a scale that looks like this:

Disagree Strongly	Disagree Somewhat	Neutral	Agree Somewhat	Agree Strongly

Click on the point in the scale that best fits how you feel about the statement. As soon as you click on a point in the scale, the next statement will appear.

As before, if you don't want to answer one of the questions, you can wait for it to time out and a new one will appear.

If you have any questions about this task, please ask the experimenter now.

If not, please click the mouse button to continue.

Figure 4.3: Instructions for language attitudes questionnaire

> sounds confident to leave *is* out, it should sound even more confident to leave it out in the unusual case.
>
> Of course, this should only be true for listeners who are familiar with AAVE. Are you familiar with AAVE?

The phrasing of the last question was fortuitous. I chose to ask about familiarity rather than actual speaking because I didn't want anyone to feel like I was accusing them of speaking a non-standard variety and because in the first exit interview I had one African-American and one non-African-American participant. By asking the question this way, it was applicable to both of them. As it turns out, some African-American participants responded to this question by saying something like "Yeah, I'm familiar with it. We talk that way at home." Other African-American participants said, "I don't talk like that, but I hear it from ... " One man said he "winced" at the preacher's bad grammar at his church. All of the African-American participants were at least familiar with AAVE. Also, some of the non-African-American participants said they were familiar with AAVE, most commonly from Hip Hop culture or from going to high schools with large African-American populations. The differences between these three groups are very interesting and will be highlighted in the discussion below.

Difficulties and missing data

Some technical difficulties and other problems led to a small amount of missing data, as described in this section.

First, a speaker setting on one computer combined with a different sampling rate for one of the test talkers caused the stimuli for that talker to not be played (i.e., the computer would show "Loading speaker..." followed by silence followed by the first scale) or to cut off in mid-sentence. When this was discovered, the program was altered so as not to try to play that talker's stimuli at all. Somewhat later, the speaker setting on the computer was fixed and the program returned to its original state. All together, seven listeners got the incomplete or silent stimuli. It was possible to tell from the output of the program which ones these were, and responses to the scales for those stimuli were discarded. Three additional listeners were only presented with stimuli from seven talkers. None of these listeners are African American, but some were familiar with AAVE nonetheless. For perhaps related reasons, the computer simply crashed part way through for one of the participants who got the incomplete or silent stimuli (Listener U). Whereas the first problem led to data being systematically missing for one test talker, the crash part way through meant that there is data missing for this listener across talkers.

In addition, some listeners didn't answer specific scales during the experiment, and the scales timed out. It's not clear whether the listeners deliberately decided they didn't want to answer these questions, or if they just hesitated too long and had the scale time out. However, the response times for the next scale in each case are long enough (at least 5 seconds) that it's clear they didn't try to click just as the scale was changing. 8 data points are missing for this reason, distributed among 7

listeners.

Further, the time out on the language attitudes questions was set at too short a setting at first. This led to a few cases where the point on the response scale that the listener chose was recorded as their response to the next question. This was evident because the response times recorded for the following question in these cases were very short, and the results were corrected accordingly.

Although the consent form stated that "you have the right to withdraw your participation at any time without penalty" and I reiterated this to the participants after they signed the forms, only one person (Listener L, an African-American participant) opted out part way through. Like the listener for whom the computer crashed, there is data missing for this listener across different talkers. She was willing to do the language attitudes questionnaire. Although I did not get a chance to do the complete debriefing with her, I did find out that she does not speak AAVE. Also, the responses she did give fit in most closely with those of Group II.

The two speakers who did not complete the main experimental task were unable to do the language attitudes questionnaire directly on the computer. Instead, I used the computer to show them the questions, but had them report their answers to me out loud. In one other case, a listener told me that he had clicked on the wrong answer for one of the language attitude questions and I recorded his corrected response.

In a couple of cases, listeners asked what the word "AAVE" meant when they saw it in the first language attitudes question. The other listeners were apparently satisfied by the definition given in the instructions for this part of the experiment.

Finally, because I had originally expected to only be recruiting AAVE speakers (and hence African-American participants), the original demographic questionnaire did not contain a question about ethnicity. This was fixed for the second day of testing, but I did not get ethnic self-identifications from the first six participants.

4.5 Results

This section presents the results of the experience. Section 4.5.1 briefly reports the results of the language attitudes questionnaire, which turned out not to correlate with responses on the judgment task. Sections 4.5.2 and 4.5.3 report the results of the judgment task and show that they confirm both parts of the hypothesis. The hypothesis is repeated here:

I Copula absence/presence in AAVE is associated with some social value.

II Copula absence/presence in AAVE is more strongly associated with that social value the more marked the environment is for each variant.

The results of the judgment task show that most listeners associated copula absence with some social value. That is, for some scales, their responses were sensitive to the presence or absence of the copula. For the two African-American groups (Groups I and II), there is a further effect of the grammatical environment, such that marked cases were rated more to the extremes of the scales. This effect is not present for Groups III-V.

4.5.1 Language attitudes questionnaire

The responses to the language attitudes questionnaire did not strongly correlate with differences between the groups or with responses on the experimental task. In fact, the most interesting language attitudes information did not come from the questionnaire, but rather from the debriefing: The self-identification as an AAVE speaker or non-AAVE speaker was a far stronger predictor of responses on the experimental task than any of the language attitudes questions. (See §4.5.3 below.) A summary of the responses to the language attitudes questionnaire are given in Table 4.14 for completeness.[24] Table 4.6 is repeated here as Table 4.13 for reference.

[24]The total number of responses for each question does not always add up to the number of listeners in a group as some listeners did not answer all of the questions.

S1	There are some situations in which it is better to speak AAVE than Standard English.
S2	A Black person who doesn't speak AAVE isn't really Black.
S3	AAVE is bad English and should never be used in any situation.
S4	It would be weird for me to speak Standard English instead of AAVE with my Black friends.
S5	When I hear a stranger speaking AAVE, I assume they are less educated.

Table 4.13: Statements for language attitudes questionnaire (reprise)

Group	statement	agree strongly	agree somewhat	neutral	disagree strongly	disagree somewhat	avg[a]
I	S1	2	3	1	2		3.62
	S2				3	8	1.27
	S3				6	5	1.55
	S4	1	1	3	2	3	2.50
	S5		3	1		6	2.10
II	S1	1	1	1	2		3.20
	S2			1		4	1.40
	S3			1	2	2	1.80
	S4			1		4	1.40
	S5	1	1			3	2.40
III	S1	3	1		1	1	3.67
	S2					6	1.00
	S3				2	4	1.33
	S4		1	2	1	2	2.33
	S5	1	4		1		3.83
IV	S1	3	1		1		4.20
	S2				2	4	1.33
	S3		1	1	1	3	2.00
	S4		1	2	1	2	2.33
	S5	1	2		1	2	2.83
V	S1		4	1		1	3.33
	S2		1	1	3	2	2.14
	S3			3	2	2	2.14
	S4		1	1	3	2	2.14
	S5	1	1	4	1		3.23

[a]Average response calculated by assigning 'agree strongly' a value of 5, 'agree somewhat' a value of 4, etc.

Table 4.14: Responses to language attitudes questionnaire

4.5.2 Copula presence vs. copula absence

This section presents the results relevant to Part I of the hypothesis, that copula absence/presence in AAVE is associated with some social value. In this analysis, the listeners were divided into the five groups shown in Table 4.7, according to their experience with AAVE and other varieties of English. Each listener heard each sentence four times and rated it each time on seven scales.[25] The sentences are repeated here in (15). In the labels for the sentences, P stands for copula presence, A for copula absence, V for following verb and N for following noun.

(15) a. PV: Yeah I know her. She's teachin me piano at Music World.
b. PN: Yeah I know her. She's my piano teacher at Music World.
c. AV: Yeah I know her. She teachin me piano at Music World.
d. AN: Yeah I know her. She my piano teacher at Music World.

To evaluate the hypothesis, I averaged the ratings across talkers, within listeners, sentences and scales. This gives each listener's average rating on each scale of each sentence. The four average ratings define an order of the sentences on each scale by each listener. Since there were 35 listeners and 7 scales, this gives a total of 245 such observations. Figure 4.4 gives six example observations, one for each possible order of two P sentences and two A sentences, where no sentences are rated exactly the same.[26]

To determine if a listener considered a scale relevant to the social value of copula absence, I compared their scores for the A sentences and P sentences. In the orders AAPP and PPAA, the sentences are differentiated according to copula absence/presence. If a listener's ratings placed the sentences in one of these two orders, the scale was considered relevant for that listener.

Out of 245 observations, AAPP appears 10 times, PPAA appears 115 times.[27] The remaining 120 observations found some other order (including orders not shown

[25]Some data is missing, as discussed in §4.4.3.
[26]In this figure and in the discussion that follows, I have regularized the scales so that the negative end is always on the left.
[27]These numbers include cases where both A sentences or both P sentences were rated the same.

AAPP: Group II, Listener N:

	1		2		3		4		5		6		7	
							AN	AV		PV		PN		
not conf	1		2		3		4		5		6		7	conf

APPA: Group I, Listener E:

	1		2		3		4		5		6		7	
								AN	PN	PV	AV			
not like	1		2		3		4		5		6		7	like

APAP: Group IV, Listener Y:

	1		2		3		4		5		6		7	
					AV		PN	AN			PV			
unrel	1		2		3		4		5		6		7	rel

PAAP: Group I, Listener K:

	1		2		3		4		5		6		7	
			PV				AV		AN	PN				
unrel	1		2		3		4		5		6		7	rel

PAPA: Group IV, Listener b:

	1		2		3		4		5		6		7	
					PV	AV	PN	AN						
not com	1		2		3		4		5		6		7	com

PPAA: Group III, Listener R:

	1		2		3		4		5		6		7	
		PV		PN AV							AN			
not com	1		2		3		4		5		6		7	com

Figure 4.4: Sample observations with no sentences rated exactly the same

in Figure 4.4 where one or more A sentences were rated the same as one or more P sentences).

Table 4.15 summarizes the 10 cases where the A sentences were both rated strictly higher than the P sentences (i.e., the 10 PPAA cases). In this table, the letters in parentheses are the listener codes. They show, for example, that only Group I's listener K rated both A sentences higher than both P sentences on two scales ('confident' and 'job').[28]

Scale	Group				
	I n=11	II n=5	III n=6	IV n=6	V n=7
comical	0%	0%	33% (R,T)	17% (a)	14% (h)
confident	9% (K)	0%	0%	17% (b)	14% (f)
educated	0%	0%	0%	0%	14% (i)
job	9% (K)	0%	0%	0%	0%
likeable	0%	0%	0%	0%	0%
polite	9% (J)	0%	0%	0%	0%
reliable	0%	0%	0%	0%	0%

Table 4.15: A* > P*, by group and scale (A* stands for AN or AV, P* stands for PN or PV. % are % of listeners in each group ranking the sentences such that A* > P*.)

Table 4.15 considers only those cases where the A sentences were ranked strictly higher than the P sentences. These scales were relevant to the social value of copula absence for these listeners. However, a slightly broader definition of relevance would allow one A sentence to overlap with one P sentence, as in (a) in Figure 4.5. In this ordering, the A and P sentences are not strictly distinguished, but the A sentences tend toward the higher end of the scale, and no A sentence intervenes between the P sentences or vice versa. Allowing observations that match the pattern in (a) of Figure 4.5 (but not those in (b) and (c) of the figure), adds four more listener/scale pairs to Table 4.15. This is reflected in Table 4.16. Although Table 4.16 remains sparse, it is interesting that the scale 'comical' produced the most PPAA orderings.

[28] In the caption of this and similar tables, A* stands for AN and AV and P* stands for PN and PV.

(a): Group I, Listener I:

AN
PV PN AV
not . . . . com
com 1 2 3 4 5 6 7

(b): Group I, Listener J:

PN
AN
PV AV
not . . . . com
com 1 2 3 4 5 6 7

(c): Group I, Listener E:

AV
PN AN PV
impol . . . . pol
1 2 3 4 5 6 7

Figure 4.5: Sample observations with some sentences rated exactly the same

Scale	Group				
	I n=11	II n=5	III n=6	IV n=6	V n=7
comical	9% (I)	20% (P)	33% (R,T)	17% (a)	14% (h)
confident	9% (K)	0%	0%	17% (b)	14% (f)
educated	9% (K)	0%	0%	0%	14% (i)
job	9% (K)	0%	0%	0%	0%
likeable	0%	0%	0%	0%	0%
polite	9% (I)	0%	0%	0%	0%
reliable	0%	0%	0%	0%	14% (i)

Table 4.16: A* '≥' P*, by group and scale

Since copula absence was rated strictly higher than copula presence in only 10 cases (and higher in only 14 cases), I will focus in the remainder of this analysis on the cases where copula presence was rated higher than copula absence. Table 4.17 gives the percentage of listeners in each group who rated the P sentences strictly greater than the A sentences on that scale (order AAPP in Figure 4.4).

Scale	Group				
	I n=11	II n=5	III n=6	IV n=6	V n=7
comical	18%	40%	17%	17%	29%
confident	27%	60%	17%	50%	14%
educated	64%	80%	100%	67%	43%
job	45%	100%	67%	67%	29%
likeable	45%	40%	17%	67%	29%
polite	27%	100%	50%	50%	29%
reliable	55%	100%	67%	50%	14%

Table 4.17: P* > A*, by group and scale

From Table 4.17, one can conclude the following: Group II listeners were most likely to judge a scale to be relevant to the social value of copula absence, Group V listeners were the least likely to do so. The scale 'educated' was relevant for the most speakers, across groups, while the scale 'comical' was the least relevant.

Tables 4.18 to 4.22 give the breakdown by listener in each group. In these tables, an X indicates that the listener rated the P sentences strictly higher than the A sentences. An / indicates that the listener rated the P sentences higher than the A sentences, but not strictly higher. (That is, the listener produced an order that is the mirror image of (a) in Figure 4.5.)

Starting with Group I (the AAVE speakers), Table 4.18 shows that while there wasn't absolute agreement as to which scales were relevant to copula absence, the scales 'job' and 'educated' were selected by most of the listeners in this group. The table also shows wide variation in the number of scales that a listener found relevant, with listener A selecting all 7 scales and listeners J and K selecting none.[29]

	Listener											
Scale	A	B	C	D	E	F	G	H	I	J	K	Total
job	X	X	X		X	X	/	/	/			8
educated	X	X	X	X	X		X	X				7
reliable	X	X	X	X	X		X					6
likeable	X	X	X	X		X						5
polite	X		X	X								3
confident	X			X		X						3
comical	X	X										2
Total	7	5	5	5	3	3	3	2	1	0	0	34

Table 4.18: P* > A* by listener and scale, Group I

In contrast with the results of Lambert et al.'s (1975) study, there is scant evidence for a positive evaluation of copula absence among AAVE speakers. While it was to be expected that AAVE speakers should rate the talkers in the copula presence condition as sounding more educated and like they have better jobs, the same isn't necessarily true for the other scales.[30] As mentioned above, the scale 'reliable' was included in the hopes that it would index in-groupness. It patterned with 'job' and 'educated' instead. One possible explanation for this is that the participants interpreted 'reliable'

[29] Note, however, that listener K rated the A sentences above the P sentences on three scales (cf. Table 4.16 above).

[30] In fact, Listener K provides an exception to this result, selecting 'job' and 'education' with the other PPAA.

as having to do with corporate reliability—i.e., reliable in the eyes of one's employer. However, these data alone do not support any firm conclusions. One might also expect listeners to score the more vernacular sentences higher on the scales 'likeable' and 'confident', but this didn't happen. Although fewer listeners produced the AAPP order for these scales than for 'job' or 'educated', only one (Listener K) produced the other order (PPAA) for 'confident' and none did for 'likeable'.

It is possible that these results reflect 'linguistic insecurity' (Labov 1966), the stigmatization of AAVE features by AAVE speakers. However, it is also possible that they reflect the influence of the experimental environment. The experiment was carried out at the community college the participants were attending, by a white experimenter and with a computer. The listeners in Group I may have been evaluating the sentences they heard with respect to the norms of the wider community.

Table 4.19 shows that Group II (African Americans familiar with AAVE, non-AAVE speakers) is much more consistent than Group I: All five speakers selected the scales 'job', 'reliable', and 'polite' as relevant, rating the P sentences strictly higher on those scales than the A sentences. Over half of the listeners in this group selected each of the scales 'confident', 'educated' and 'likeable' as relevant.

	Listener					
Scale	L	M	N	O	P	
job	X	X	X	X	X	5
reliable	X	X	X	X	X	5
polite	X	X	X	X	X	5
confident	X	X	X		/	4
educated	X	X		X	X	4
likeable	X	X		X		3
comical	X		X			2
Total	7	6	5	5	5	

Table 4.19: P* > A* by listener and scale, Group II

The listeners in Group II all indicated that they were familiar with AAVE, but then distanced themselves from the variety. They did this by saying things like "I wince whenever my preacher splits his infinitives" or "I'm trying to teach my daughter to speak Standard English, but she's picked up AAVE at school." It seems that these speakers may have decided at some point to use Standard English rather than AAVE, although both may have been available to them. Having made such a decision, they would probably have negative associations with AAVE features, and such features would probably be very salient for them.

As shown in Table 4.20, while Group III (non-African Americans familiar with AAVE, non-AAVE speakers) selected fewer scales than Group II, there is still complete agreement on one scale ('educated'), and fairly good agreement on one other ('polite'). As with Group I, there is rather wide variation in how many scales a given listener selected.

	Listener						
Scale	Q	R	S	T	U	V	
educated	X	X	X	X	X	X	6
polite	X	X	X	/	/		5
job	X	X	X	X			4
reliable	X	X	X	X			4
likeable		X				/	2
confident	X						1
comical	X						1
Total	6	5	4	4	2	2	

Table 4.20: P* > A* by listener and scale, Group III

Table 4.21 gives the results for Group IV (native speakers of English, no familiarity with AAVE). In this group, 'job', 'educated' and 'likeable' are selected the most often, but none by all 6 listeners. There is also wide variation in how many scales a given listener selected. Listener W selected all 7, and Listener b only one, with AAPP (see Table 4.15 above).

Finally, Table 4.22 gives the results for Group V (non-native speakers of English, no familiarity with AAVE). This table is the sparsest of all, although there is some agreement on 'educated' as a relevant scale.

The general picture that emerges from these tables is as follows: Group II had the most agreement about which scales were relevant, but it's not clear that this

	Listener						
Scale	W	X	Y	Z	a	b	
job	X	X	X	X			4
educated	X	X	X	X			4
likeable	X	X		X	X		4
confident	X	X	X				3
polite	X	X	X				3
reliable	X	X		X			3
comical	X		/				2
Total	7	6	5	4	1	0	

Table 4.21: P* > A* by listener and scale, Group IV

	Listener							
Scale	c	d	e	f	g	h	i	
educated	/	X	X	/	X	/		6
likeable	X	X		/				3
job	X	X			/			3
polite	X		X					2
confident	X					/		2
comical				X			X	2
reliable			X					1
Total	5	3	3	3	2	2	1	

Table 4.22: P* > A* by listener and scale, Group V

should be interpreted as this group having a strong convention of associating a social value such as 'reliable' with copula absence. It is also possible that copula absence was simply the most salient for this group, and generally evaluated negatively. For Group III, where fewer scales were selected overall, it is more striking that the scale 'educated' was selected by all of the speakers. Group V listeners selected the fewest scales as relevant, and Groups I, III and IV fall somewhere between Groups II and V. The fact that Group I didn't select more scales as relevant, and especially the fact that they didn't rate the A sentences higher than the P sentences on all but a few scales, may have been the result of a conflict between their own attitudes towards

AAVE and the normative attitudes towards AAVE that they were sensitive to in the experimental session.

Nonetheless, the evidence presented in this section provides support for part I of the hypothesis: In the experimental task, listeners based their social evaluation of the talkers in part on the presence vs. absence of the copula. The heterogeneity in the pattern of scales selected (especially in Group I) raises intriguing further questions, but the fact that the majority of listeners selected some scales is sufficient to allow for testing of part II of the hypothesis.

4.5.3 Effect of the following grammatical environment

This section reports the results that pertain to part II of the hypothesis: that there should be an interaction between the grammatical environment and the social value of the variable. In particular, the hypothesis states that copula absence/presence in AAVE is more strongly associated with its social value the more marked the environment is for each variant. That is, part II of the hypothesis will be confirmed, if, for any scale treated as relevant by a listener, copula absence before a noun (AN, the more marked case) is rated lower than copula absence before a verb (AV, the less marked case) or copula presence before a verb (PV, marked case) is rated above copula presence before a noun (PN, marked case). If a talker sounded relatively uneducated to a listener in the copula absence condition, they should have sounded especially uneducated when using copula absence before a noun. To put it yet differently, to the extent that a disfavorable social value is associated with copula absence, that disfavorable social value should be heightened in the AN sentence with respect to the AV sentence. Conversely, to the extent that a favorable social value is associated with copula presence, that favorable social value should be heightened in the PV sentence with respect to the PN sentence. Such a result would show that speakers have knowledge of the non-categorical constraint.

Returning to the ordering of the sentences on the scales illustrated in Figures 4.4 and 4.5, and only considering the AAPP and PPAA orders, there are two orderings that conform to the hypothesis:

(16) a. AN < AV < PN < PV
  b. PV < PN < AV < AN

In these orders, the two marked cases (PV and AN) are at the ends. Six other orders (given in (17), contradict the hypothesis:

(17) a. AN < AV < PV < PN
  b. AV < AN < PN < PV
  c. AV < AN < PV < PN
  d. PV < PN < AN < AV
  e. PN < PV < AV < AN
  f. PN < PV < AN < AV

However, as noted above, there were only 10 listener/scale pairs in which both A cases were rated above both P cases. Therefore, we will focus on the P > A type, i.e., the orders (16a) and (17a-c).

The complete orderings in (16) assume that both copula presence and copula absence are socially significant, and thus should have social values sensitive to the following grammatical environment. However, it could be that only one variant (absence or presence) in fact bears social value for a given population. That is, copula absence could be socially and stylistically unmarked, available for use at any time, while copula presence is socially and stylistically marked (or vice versa).[31] If copula absence is unmarked, then the ratings for the copula absence sentences should reflect a baseline, while the ratings for the copula presence sentences reflect the additional effect of copula presence. In this case, one would only expect to see an effect of the following grammatical environment on the evaluation of copula presence. Of course, it is not possible to tell a priori which variant (or both) is socially meaningful, and different groups may indeed differ on this point. Therefore, the effect of the following grammatical environment will be considered separately for copula presence and copula absence.

[31] This is simplifying things a bit, as what is socially and stylistically marked surely varies with the situation and the interlocutor.

If we consider copula presence and copula absence separately, the hypothesis-confirming order (16a) breaks down into two orders, given in (18).

(18) a. AV/AN < PN < PV
b. AN < AV < PV/PN

These orders specify that both P sentences must be rated higher than both A sentences, but the relative ranking of the N and V sentences is only relevant for either P or A. In this form, the disconfirming orders would be:

(19) a. AV/AN < PV < PN
b. AV < AN < PV/PN

The null hypothesis is that the following grammatical environment has no effect, so the orderings of the V and N sentences should be equally distributed between (18) and (19). The alternative hypothesis is that the following grammatical environment systematically affects the ratings, and one ordering of the V and N sentences occurs significantly more than chance. In particular, the ordering in (18) should come up significantly more than chance.

The Exact Binomial Test can distinguish between the null and alternative hypotheses in a dataset like this one. The results of this test are given in Table 4.23. The first column of this table gives the group. The second column gives the total number of observations for the group. For example, there were 11 listeners in Group I and 7 scales, giving 77 observations for Group I. The third column gives the number of observations in which the listener rated both of the P sentences higher than both of the A sentences on the scale. Since this part of the hypothesis concerns the effect of the following grammatical environment on the social value of the variable, it was only tested within those cases where the listener found the scale relevant to that social value. (For Group I, there were 31 such cases.)

The fourth column gives the proportion of observations that matched the predicted order of AV and AN. The denominator in the fourth column is less than the number in the third column, as any cases where AV averaged the same as AN were discarded.[32]

[32] That is, they were considered to be cases where the instrument was not sensitive enough to tell which way the order went.

Group	Total N	P* > A*	AV > AN	p	PV > PN	p
I	77	31	[a]12/28	0.2888	*18/23*	*0.0053*
II	35	26	*23/25*	*0.0000*	16/25	0.1148
III	42	20	[a]6/17	0.1662	9/18	0.5927
IV	42	22	[a]4/14	0.0898	8/14	0.3953
V	49	16	5/9	0.5000	[b]3/10	0.1719
I&II	112	57	*35/53*	*0.0135*	*34/48*	*0.0028*
III-V	133	58	[a]15/40	0.0769	[b]20/42	0.4388

[a] AN > AV more frequent.
[b] PN > PV more frequent.

Table 4.23: Effect of following grammatical environment, all groups

The next column gives the p values produced by comparing this proportion to 1/2 (chance distribution) with the Exact Binomial Test. The only significant effect of the following grammatical environment on copula absence was for Group II, and it goes in the direction predicted. For Groups I and III-V, the distribution of the two orders of AN and AV is not distinguishable from chance. The last two columns give the results for the effect of the following grammatical environment on copula presence. In this case, the only significant effect was for Group I, again in the predicted direction. Finally, in order to make sure that the lack of significant results for Groups III-V was not due to the small sample size, the last row of the table gives the results for the combined group III-V. Even for this larger group, no significant result emerges.

These results show that, for Groups I and II only, there was significant effect of the grammatical environment on the social evaluation of forms of the copula. When listeners in Group I rated copula presence higher (on some scale) than copula absence, they also rated copula presence before a verb higher than copula presence before a noun significantly more often than chance. This means that the social value of copula presence for these listeners is intensified in the marked environment. There are two possible ways that this could come about. The first is that these listeners know that a following verb is a marked environment for copula presence, and judge that a speaker would only use copula presence in that environment if s/he were particularly emphatic about expressing the social value of copula presence. The second is that the social value of the copula presence is encoded in the grammar as being dependent on the part of speech of the predicative phrase. In this case, the fact that copula presence is rarer before verbs than before nouns would follow from speakers avoiding copula presence before verbs except when they really wanted to express that more intense social value. Either way, the listeners have some direct, if tacit, knowledge of the non-categorical constraint. Similar remarks hold for Group II with respect to copula absence.

Why should there be a difference between Groups I and II? The result that Group I listeners only showed an interaction with the grammatical environment in copula presence would follow if copula absence is socially and stylistically unmarked for Group I, while copula presence is socially meaningful. Similarly, the results for Group II would follow if it is copula presence that is unmarked and copula absence that is marked for this group. Recall that Group I listeners self-identified as AAVE speakers, while Group II listeners said they were familiar with AAVE but did not self-identify as AAVE speakers, and indeed distanced themselves from AAVE. It would seem quite plausible for the latter group to find copula presence unmarked and copula absence as meaningful (and indeed, as indicative of various negative personality traits). Copula absence would be less remarkable for the AAVE speakers (Group I). It is possible, although less predictable, that they associate social value only with copula presence. Another possibility, raised in more detail in Chapter 6, is that the fine-grained distinctions in social value according to the grammatical environment depend on the variant in question being relatively infrequent. Copula absence is obviously infrequent for non-AAVE speakers (Group II), and copula presence is more infrequent for AAVE speakers than for non-AAVE speakers. Whether it is infrequent enough, and indeed what contexts that frequency is calculated over, remains to be seen.

Although there are important differences between the results for Groups I and II, both groups' responses are sensitive to the same non-categorical constraint. One might ask why the listeners in Group II, as non-AAVE speakers, were aware of this pattern while the listeners in Group III, who also claimed to be familiar with AAVE, were not. The answer is most likely that the African Americans (Group II) had much more extensive and involved experience with AAVE as listeners than did the

listeners in Group III who were not African American. The contrast between Groups I and II, on the one hand, and Groups III-V on the other, is an important result: only speakers who are sufficiently familiar with AAVE were showed knowledge of the non-categorical constraint, although speakers in other groups did find copula absence/presence relevant to their evaluations of the talkers. This shows that the non-categorical constraint is specific to AAVE, and not something latent in the structure of English in general.

4.6 Conclusion

This experiment remains a pilot study in many respects. The results need to be confirmed for a larger population, for other non-categorical constraints and for other variables. Nonetheless, the results are suggestive, and raise the interesting possibility that speakers have knowledge of non-categorical constraints on sociolinguistic variation and that these constraints are further relevant to the social evaluation of any given instance of a variable. Chapter 5 evaluates existing theories of non-categorical constraints on sociolinguistic variation against the results of this experiment and other aspects of socially meaningful variation.

Chapter 5

Previous theoretical approaches to non-categorical constraints

5.1 Introduction

The purpose of this chapter is to review and critique previous approaches to accounting for non-categorical grammatical constraints on variation. Three such approaches will be considered: The variable rule approach of Labov and others (§5.2), the Optimality Theoretic approaches of Reynolds and Nagy (1994), Anttila (1997) and Boersma and Hayes (1999) (§5.3), and the functionalist approach of Kiparsky (1972, 1988) (§5.4). This first section provides some background for the discussion, in three parts: §5.1.1 discusses the problems that sociolinguistic variation raises for theories of grammar, from Weinreich et al.'s (1968) challenge to the idea of grammar as a homogeneous system to the problem of non-categorical grammatical constraints, and finally the association of social value with variation. This last issue is taken up in more detail in §5.1.2. §5.1.3 outlines the data to be accounted for by any analysis of socially meaningful, grammatically constrained variation against which the three previous accounts considered will be judged.

5.1.1 Variation and grammar

According to Weinreich et al. (1968), the field of linguistics, from the work of Hermann Paul in the late 1800s through the generative phonology of the 1960s, had been mistakenly searching for homogeneity in language, on the basis of the misguided assumption that only homogeneous systems can be structured. They argue that this misconception was behind Saussure's paradox: how do slow gradual changes in language use create (new) homogeneous language systems? If language systems are not homogeneous, but rather variable, the paradox dissipates. Another example of the effect that the assumption of homogeneity had on linguistic theory is Bloch's now infamous definition of *idiolect*:

> The totality of the possible utterances of one speaker at one time in using a language to interact with one other speaker is an *idiolect*. ... The phrase 'with one other speaker' is intended to exclude the possibility that an idiolect might embrace more than one STYLE of speaking: it is at least unlikely that a given speaker will use two or more different styles in addressing a single person. (Bloch 1948:7, §1.7)

In trying to define variation out of the purview of linguistics, Bloch arrived at an object of study that was too narrow—precluding, for example, a theory of how idiolects relate to each other, within or across speakers—and, at the same time, not narrow enough, for individual speakers can style-shift even while talking to an individual hearer.

Fries and Pike (1949) take the first step away from the assumption of homogeneity, stating that "Socially pertinent differences of style cannot safely be ignored; they must be handled in some way in our phonemic assumptions and procedures" (p. 29). Fries and Pike's proposal is to allow for coexistent phonemic systems when some phenomenon doesn't fit the general phonological pattern abstracted from a language. The case they explore in most detail is the problem of new phonemes or contrasts brought into a language by loanwords. However, they also suggest that the same method would do in the case of stylistic differences in the speech of one speaker and in the case of "a conflict in the system of sounds of a single speaker" during a time of change in the phonemic system of a language.

However, the postulation of coexistent phonemic systems can also be seen as a last-ditch effort to preserve the homogeneity of linguistic systems in the face of variable data that Fries and Pike realized require honest treatment. Weinreich et al. (1968) propose instead to abandon the assumption that structure is only to be found in homogeneity, and argue that a realistic description of any synchronic stage of a language as well as an empirical theory of language change are only possible if we recognize that language can be (and is) both heterogeneous and structured.

Weinreich et al. establish that phonological variation is structured in two respects: at the level of the linguistic system and at the level of the speech community. Phonological variation is structured at the level of the linguistic system in that variation in one part of the grammar is associated with variation in another. For example, Weinreich et al. find that the (variable) raising and backing of (ah) in *father* and raising of (oh) in *coffee* in New York are related: the more a speaker or group of speakers does one, the more they do the other. This fact has a structural explanation: if (ah) were raised and backed while (oh) stayed in place, the distribution of the phonemes would overlap and distinctions would potentially be lost.[1]

Phonological variation is structured at the level of the speech community because the hierarchical organization of speech styles parallels the (hierarchical) socio-economic organization of the community. This too is true in two respects: First, in Labov's (1966) study of a number of variables in New York City, he found that the higher a speaker's socio-economic status, the lower their use of stigmatized variants. Second, in the same study, Labov also found that speakers used more stigmatized variants in more casual speech and fewer stigmatized variants in more formal, or careful, speech, reproducing the community's hierarchy of speech styles within their own repertoire, albeit on a smaller scale. Although this interpretation of the relationship between variation, the social structure of a community, and style has been challenged (see §5.1.2), Weinreich et al.'s (1968) main point still stands: variation

[1] Bender (1998) provides a syntactic example: Avoiding the use of the epicene pronoun *he* when the gender of the referent is unknown leads to a reduced use of singular definite generic NPs with human reference (e.g., *the poet* in *The poet chooses his words with care.*) in the texts studied. Here the explanation concerns linguistic structure in a somewhat different sense: singular definite generic NPs introduce precisely the kind of referent that requires an epicene pronoun.

is not merely random fluctuations to be assumed away, but instead indicative of the interface between language and social structure.

The other important point established by Weinreich et al. is that variation cannot be eliminated by narrowing the object of study down to one style of one speaker, etc. Variation is inherent to the language of communities and to the language of individuals. What differentiates styles (within or across speakers) is not the presence or absence of given variants, but the relative frequency of occurrence.

The preceding arguments all concern phonology, but what of syntax? The first question is whether sociolinguistic variation exists at the level of syntax in the same way that it exists in phonology. This, in turn, breaks down into two questions: Is there such a thing as a syntactic variable? And, if so, do syntactic variables interact with social structure in the same way phonological variables do?

A variable, standardly, is a choice-point in the linguistic system that allows for two (or more) ways of saying the same thing. The variable (ah) mentioned above groups together multiple pronunciations of the first vowel in *father*. The choice of which variant is pronounced has no (denotational) semantic effect. As Lavandera (1978) points out, this cannot be straightforwardly extended to syntax, where typically a choice-point in syntax involves a choice between different morphemes. Since morphemes, unlike phones, are typically meaningful, many choice-points in syntax do not provide multiple ways of saying the same thing. Rather, one is faced with multiple utterance-types that do not mean exactly the same thing, although they may be semantically related.

However, not all morphemes necessarily carry referential meaning. Function words and inflectional affixes are standardly analyzed as semantically vacuous formal elements. So variation concerning such elements should in fact amount to multiple ways of saying the same thing. Examples include variable negative concord (1), variable subject-verb agreement (2),[2] and, to some extent, variable copula absence. (I hedge there because the copula does carry tense information, although, as discussed in Chapter 3, the alternative construction may carry the same tense information.)

[2] From Guy 1996:238.

(1) a. I didn't see anything.
b. I didn't see nothing.

(2) Popular Brazilian Portuguese:

a. Eles fizeram uma bagunça
They made.pl a mess

b. Eles fez uma bagunça
They made.sg a mess

Likewise, variables having to do with the order of constituents are potentially analyzable as different ways of saying the same thing. One example is the order of the particle and the complement in English verb-particle constructions:

(3) a. I took the garbage out.
b. I took out the garbage.

Alongside such variables are other choice-points that cannot be analyzed as different ways of saying the same thing. For instance, an agentless passive sentence differs from its active counterpart in leaving the agent unexpressed (4).

(4) a. Eight talkers were recruited to record the sentences.
b. I recruited eight talkers to record the sentences.

Although it is often clear from context who the agent is, there remains a clear semantic difference.[3] Sentences such as (4a) are characteristic of English scientific discourse (e.g., in descriptions of experiments). One possible explanation for this is that, by backgrounding the agent, they make it sound like the experiment was carried out perfectly, with no human agent to potentially do things wrong. As reproduceability of results is central to modern scientific practice, backgrounding the agent makes writing sound scientific.

[3] Weiner and Labov (1983), in their variationist study of the passive, contrast agentless passives with active sentences with subjects like generic *they* that give little or no information on the identity of the agent. In this case, the active and the agentless passive are closer to being two different ways of saying the same thing. However, the fact remains that the choice between active and passive is available even when they are patently not two different ways of saying the same thing.

Lavandera (1978) argues that sociolinguistic variables were originally restricted to alternate ways of saying the same thing, not for any sound theoretical reason, but out of fear that finding that different groups of people say different things would provide evidence that could be used to support prejudices against one or more of the groups studied. She suggests that while the intentions behind this fear are good ones, the overall social purpose would be better served by practicing more theoretically sound science and remaining vigilant for any abuses of the results. On her view, the notion of sociolinguistic variable should be applied to any set of linguistic forms provided that (a) the forms share some kind of functional equivalence,[4] and (b) the relative frequencies of the forms (and not just the forms themselves) carry some social import.

The question of whether the social value of a variable is in the individual instances or in the pattern will be taken up again in Section 5.1.3 below. For now, I would like to ask if any syntactic variables vary with social factors in the same way that phonological variables do. Although syntactic variables have not been as well studied as phonological variables, some have been shown to correlate with social factors. Cedergren and Sankoff (1974) found that *que*-deletion in Montreal French correlates with social class. Eckert (2000) found that rates of negative concord among the Detroit adolescents in her study correlated with various socioeconomic facts about the speakers' parents (most notably, the mother's level of education) and with social categories within the high school. In fact, Chambers (1995:51) claims that "[g]rammatical variables tend to mark social stratification more sharply [than phonological variables]."

However, despite claims to the contrary by Mufwene (1992), the mere fact of variation does not pose the same problems for modern theories of syntax as it did for structuralist phonology. Mufwene (1992:234) argues that there is in syntactic theory a "mistaken working assumption that grammars are monolithic institutions with rules that are consistent with each other," citing AAVE copula absence as one example. He proposes that copulaless sentences should be analyzed in terms of an alternative rewrite rule for the category S:

(5) S $\longrightarrow$ NP PredP

[4] Lavandera leaves the precise definition of this notion of functional equivalence to later work.

As shown in Chapter 3, a sign-based view of grammar does not require that there be only one way to build an S.[5] In fact, the existence of multiple rewrite rules for the same category is fundamental to a construction-based approach. Even in GB, where there appears to be general assumption that the phrase-structural backbone is the same across languages, or in Minimalist syntax where different derivations compete with each other for grammaticality, variation in syntax can be accommodated by deriving the different variants from underlying forms with different (possibly phonologically null) lexical items. The one exception to this kind of compatibility with intraspeaker variation is OT syntax, to the extent that it posits semantic structures as underlying (or input) forms. However, as discussed in §5.3, the classical OT model has been extended in various ways to accommodate variation in phonology, and all of these extension can be applied to OT syntax.

However, while optionality does not pose any problem, probabilistic or frequentistic application of rules does. Variable rules (see §5.2 below) were proposed by Labov (1969) and Cedergren and Sankoff (1974) as an extension of generative grammar. In particular, variable rules were presented as a more precise version of optional rules. Optional rules applied some of the time, while variable rules applied some specific percentage of the time. Generative grammarians (including Butters 1972 and Kay and McDaniel 1979) responded by pointing out that this was a misinterpretation of generative grammar: a generative grammar produces sentence TYPES, and an optional rule adds a sentence type to the set of sentences types generated by the grammar. Frequency of occurence is a matter of TOKENS. To adorn generative rules with probabilities to produce such frequencies of occurrence is to misconstrue them as rules of performance—i.e., rules that people apply when producing a sentence to say. Thus the general treatment by generative grammarians of frequentistic data is to rule such effects as outside the grammar proper, i.e., not a matter of competence but of performance.

While some frequencies are strictly epiphenomenal, (cf. Chomsky's observation about the relative frequencies of *I live in New York* and *I live in Dayton, Ohio*[6])

[5] Although this does not turn out to be the best analysis of copula absence in AAVE.
[6] From a lecture to the 1964 LSA summer institute, cited in Halliday 1991:30.

this is not so clearly the case for the frequentistic effects in sociolinguistic variation. In particular, there is the problem of non-categorical grammatical constraints on variation.[7] As discussed in the previous chapter, the rate of copula absence in AAVE is sensitive to the part of speech of the predicate. While it is possible that these effects are a matter of (universal, functional) performance factors (see §5.4), the results of the experiment reported in Chapter 4 indicate that they are also a matter of linguistic knowledge. I will suggest in Chapter 6 that they may also be a matter of linguistic competence. Here what is relevant is that if any frequentistic or probabilistic effects are found to be a part of linguistic competence, such a result would be incompatible with mainstream generative grammar.

The final problem that sociolinguistic variation raises for mainstream, modern theories of syntax is its intimate connection with social meaning. In mainstream theories (and here I mean GB and Minimalism), linguistic competence is supposed to be relegated to a separate module of the mind, and syntax to a separate module within linguistic competence. If syntactic variation can carry social meaning, then there must be some way for grammar and social meaning to interface. Further, the results of the experiment reported in Chapter 4 show that social meaning is not simply a property of individual variants, but is also intertwined with the syntactic context of the variants. This renders any theory in which social meaning is 'read off' the syntax outside the grammar less plausible, as such a theory would have to reproduce non-trivial syntactic structure outside the grammar. Section 5.1.3 outlines the data a theory of non-categorical constraints must account for (whether it's a theory of competence or a theory of performance). First, however, the next subsection discusses social meaning and the social value of linguistic variation.

[7] Cf. Cedergren and Sankoff 1974, 333:

> The notion of optionality [of rules of grammar–EB] fails to capture the nature of the systematic variation which exists even on the level of the grammar of a single individual. It does not permit the incorporation of relativity or covariation between the presence of certain features in the linguistic environment of a rule and the frequency of operation of the rule.

5.1.2 Social meaning

Weinreich et al. (1968) find the connection between diachronic and synchronic perspectives on language in a model of language that is heterogeneous and intimately connected with the structure of society. They write:

> [The empirical studies reviewed] have confirmed the model of an orderly heterogeneous system in which the choice between linguistic alternants carries out social and stylistic functions, a system which changes with accompanying changes in social structure. (1968:162)

Labov (1963) fleshes this out with a particular example: centralization of the nucleus of the diphthongs (ay) and (aw) on Martha's Vineyard. In that community at that time, high rates of centralization of the nuclei (traditional Vineyard pronunciation) were associated with resistance to the change to a tourism-based economy. Low rates of centralization (i.e., approximation to the mainland standard) were associated with enthusiasm for the economic changes. Further,

> Centralized speech forms are then a part of the dramatized island character which the Chilmarker assumes, in which he imitates a similar but weaker tendency in the older generation.

Labov speaks of this in terms of younger speakers identifying with some older speakers as a reference group and then hypercorrecting their speech in an attempt to be like the reference group. In current terms (e.g., those of Eckert and McConnell-Ginet 1992), centralization of the nucleus of (ay) and (aw) is one symbolic resource used in the on-going construction of the "dramatized island character", that is, the old Yankee fisherman to whom (according to one local ideology) Martha's Vineyard "rightly" belongs.

As variationist studies expanded to include larger samples of larger speech communities, the treatment of social factors was simplified to seeing them as simply a matter of social "address": the use of certain variants or certain patterns of use are markers of certain pre-established and global groups (such as "women", "lower middle class", "African-American").[8] On this view, a change travels through the social

[8]Compare the old Yankee fisherman character. Not only was this particular social category local to Martha's Vineyard, it also was constituted by a range of particular activities, attitudes and ways

structure and is reflected in the speech of people from different social categories according to whether the change has reached them. When they are not paying much attention, people speak their vernacular: the system they learned growing up, the one appropriate to their social address. When paying attention, they may move away from that vernacular, typically correcting their speech towards some prestige variety. This notion of vernacular is implicit in Labov's repeated claim that the vernacular is the most systematic speech style.

The social-address view of variation has been challenged on many dimensions. Most important here are the following: The view of social categories as pre-existing (nearly "natural") categories and the view of language as merely reflecting the social categories a person belongs to. The proposed alternative is a view of social categories as socially constructed and language as one tool used in that process of construction. Eckert and McConnell-Ginet put it succinctly:

> It is the mutual engagement of human agents in a wide range of activities that creates, sustains, challenges, and sometimes changes society and its institutions, including gender and language. (1992:462)

Crucial to this view is the notion that institutions or social categories such as gender are highly textured, or contentful. It is often said that gender differentiation is the social elaboration of a rather simple biological differentiation.[9] What this means is that social categories such as "male" (or "masculine") and "female" (or "feminine") are reifications of a set of properties, practices, or stances.[10] Ochs (1992) gives the example of two sentence-final particles in Japanese. The particle *ze* expresses coarse intensity and the particle *wa* expresses delicate intensity. The properties coarse and delicate are partially constitutive of the categories masculine and feminine, respectively. One thing about women, according to Japanese ideology, is that they are delicate, even when expressing intensity. One thing about men, according to the same ideology, is that they are (or can be) coarse about things such as expressing

of being.

[9]Although the widely held assumption that the biological differentiation at least is clear-cut has also been challenged. See, for example, Epstein 1990.

[10]This applies equally to local social categories such as the old Yankee fisherman on Martha's Vineyard or "jocks" and "burnouts" among the students of Belten High studied by Eckert 2000.

intensity. To use the particle *ze* is to be coarse, or in Ochs's terms to directly index coarse intensity. The use of the particle thus indirectly indexes gender: used by a man, it indexes masculinity. Used by a woman (against expectations) it probably indexes something more like anti-femininity. By the act of speaking one of these particles, the speaker has performed an instance of coarse or delicate intensity, which is part of being masculine, feminine, anti-masculine or anti-feminine, depending on what was said, by whom.[11]

Thus social categories such as gender are textured and contentful. They are also crucially continually reproduced/reenacted. Taking for the moment the perspective of a child being socialized, it is clear that the only way to learn a category and the practices that make it up is to experience it in some way—through watching, interacting with, or perceiving the expectations of others. Social categories only exist because they have a history, and only persist from generation to generation because they are reproduced in everyday acts. For adults (socialized members of society) reified notions such as femininity in all of its details might persist for some time even if their performance was stopped. Nonetheless, they are still reproduced by any act that depends on the reification for interpretation, and can be subtly changed by acts that challenge the existing notions.

To take a concrete example, Subaru currently has a commercial out in which three female athletes (tennis player Martina Navratilova, golfer Juli Inkster, and skier Diann Roffe-Steinrotter) are all shown saying "What do I know about power/control/grip?" Then they each get into one of the Subaru SUVs and say something like "What horsepower!" The commercial ends with Martina Navratilova getting out of her car and saying, "But what do we know? We're just girls." This commercial is openly highlighting and disputing certain aspects of the gender system in the US, for example, that only men appreciate those aspects of cars. At the same time, it is constitutive of another piece of the gender system, in particular, what it means to be a female athlete. According to this commercial, athletic prowess goes along with being knowledgeable about cars and, more generally, taking on "male" characteristics.

[11] This whole picture is complicated, of course, by the possibility of alternative femininities and masculinities. See, for example, Connell 1995.

The next question is, how does language fit into this process? The answer is: ubiquitously. As noted with Ochs's example of Japanese sentence-final particles above, particular linguistic forms can index social stances that are in turn constitutive of social categories. In every use and interpretation of such forms, that aspect of the linguistic system, the stance, and the social category are reproduced. But the interpretation of socially significant linguistic forms is not capricious, but rather mediated by ideology. Irvine and Gal (2000) argue that the relationship between form and social meaning is cemented by a process of iconization:

> *Iconization* involves a transformation of the sign relationship between linguistic features (or varieties) and the social images with which they are linked. Linguistic features that index social groups or activities appear to be iconic representations of them, as if a linguistic feature somehow depicted or displayed a social group's inherent nature or essence. This process entails the attribution of cause and immediate necessity to a connection (between linguistic features and social groups) that may be only historical, contingent, or conventional. (2000:37)

From this point of view, the connection between the particles *ze* and *wa* to coarse/gentle intensity and to masculinity/femininity is non-arbitrary to speakers of Japanese. *Ze* IS coarse and *wa* IS gentle, so of course men (who are coarse) say *ze* and women (who are gentle) say *wa*.

One effect of the process of iconization is that we are often unaware of the fact that we are getting social information from language. Ochs writes:

> Competent members of every community have been socialized to interpret these meanings and can without conscious control orchestrate messages to convey social meanings. (1992:338)

That is, because of the apparent iconic relationship between form and social value, we plan to be polite rather than planning to be polite by phrasing a command as a request. Likewise, a native speaker of Japanese need only intend to express gentle intensity, not intend to express gentle intensity by using *wa*.[12] This is no different

[12] As a non-native speaker of Japanese, the experience is entirely different: In my own experience with Japanese, I've often found myself having to intentionally plan on using politeness markers and other aspects of feminine speech when I wanted the effect of those forms.

from the way in which we plan to talk about cats as opposed to planning to talk about cats by using the word *cat*.

Neither the process of iconization nor the reification of social categories implies that social meanings are fixed. California Style Collective (1993), Eckert (2000), Irvine (in press) and Coupland (in press) paint a picture of the social world as a landscape inhabited by salient people with salient characteristics. Individual speakers attempt to make their place in that landscape through a process of 'bricolage' (California Style Collective 1993): the piecing together of a personal style out of symbolic resources made meaningful by their association with other people and groups in the social landscape, ideologies about those people, and the speaker's own attitudes. According to Coupland (in press), this process of identity building is as much for the speaker's own sense of self as it is a display or performance for others. Further, identity building is not once-and-for-all, but rather highly relational. Irvine (in press) provides an example from the Wolof opposition of noble and griot[13] castes. Each caste is associated with a supposedly distinct style of speech. In practice, Irvine found a continuum of speech styles whose endpoints represented public speech by each caste. The middle is used relationally:

> Two persons who belong to one and the same caste will differentiate their speech along the same stylistic axis that differentiates castes from each other, in order to represent subtler differences of rank (such as lineage seniority), or to define an activity, such as petitioning, that is reminiscent of intercaste relations. (in press)

Note also that the social construction of identity is not necessarily entirely voluntary/in the control of the speaker. The role of the listener is important, too: if listeners do not acknowledge or do not understand the speaker, they can frustrate a speaker's efforts to build and project a particular social identity. Eckert and McConnell-Ginet discuss the concept of "symbolic privilege" or the privilege to "assume [one's] own positions to be norms toward which everyone else orients" and to "judge other positions while supposing [one's] own to be invulnerable to less privileged assessment" (1992:483).

[13] A low-ranking caste "whose specializations are public speech making, praise singing, music and the rhetorical and communicative arts in general." (Irvine 1990:132)

Summary In Searle's (1995) study of *The Construction of Social Reality*, he proposes that all social facts basically consist of people mutually agreeing to assign a certain status to a certain entity. At the basic level, statuses are assigned to physical phenomena, for example, a smooth, heavy piece of glass with patterns inside can be assigned the status of a paperweight. Statuses can be assigned recursively, so that that same heavy piece of glass *qua* paperweight can be assigned the status of belonging to someone. As a possession, it can be assigned the status of a gift, and of belonging now to someone else. Searle schematizes this status-assignment as "X counts as Y (in context C)". This concept might be useful for reviewing the relationship between language, social categories and identities explored in this section.[14]

What we have is a recursive system of status-assignment. For Japanese speakers, saying *ze* counts as expressing coarse intensity, and expressing coarse intensity (along with a variety of other acts) counts as being normatively masculine. For Eckert's (2000) Belten High speakers, using negative concord counts as (one aspect of) orienting one's style in the social landscape towards Detroit and urban culture. Orienting towards Detroit and urban culture counts as (one aspect of) being a burnout. When a linguistic form counts as something in the social landscape, we say that it has social meaning.

There is an important aspect of social meaning (and indeed other aspects of language and other social institutions) obscured by the phrases 'counts as' and 'mutual agreement' in the preceding paragraphs: the assignment of social meaning is not a process in which all speakers participate on an equal footing. Speakers may believe they agree with others on the social value of a variant, say, while the other parties in fact have another idea. Further, people in positions of power often have the privilege to assume that others share their 'mutual' understanding, since they essentially have the power to assign the blame for any misunderstanding to a less powerful interlocutor. This gives powerful people more meaning-making rights (e.g., Eckert and McConnell-Ginet 1992). This isn't unique to social meaning, but pervades language use. Mainstream English speakers do not usually attempt to learn AAVE, but rather

[14] Note, however, that where Searle speaks of mutual agreement, Ochs speaks of socialization and Irvine and Gal of iconization.

expect AAVE speakers to accommodate their (more standard) variety.

5.1.3 Data to be accounted for

Labov (1966) and Weinreich et al. (1968) find social meaning (such as it was for them) in the differing frequencies of each variant across different social groups. The discussion of social meaning in the construction of social categories and the social landscape above focussed on individual forms. Where does social meaning reside? Given a sociolinguistic variable, there is both a pattern of variation and individual instances of each variant. Eckert and McConnell-Ginet write:

> Individual agents plan and interpret situated actions and activities, but their planning and interpretation rely on a social history of negotiating coordinated interpretations and normative expectations (and in turn feed into that history). (1992:474)

Similarly, Bakhtin (1986) writes, referring to an arbitrary speaker:

> He is not, after all, the first speaker, the one who disturbs the eternal silence of the universe. And he presupposes not only the existence of the language system he is using, but also the existence of preceding utterances—his own or others'—with which his given utterance enters into one kind of relation or another (builds on them, polemicizes with them, or simply presumes that they are already known to the listener). (p. 69, translation by Vern W. McGee)

This is reminiscent of Halliday's (1991) description of the interrelation between grammar and text with reference to the relationship between climate and weather:

> To the 'instance' observer, the *system* is the potential, with its set of probabilities attached; each instance by itself is unpredictable, but the system appears constant through time. To the 'system' observer, each *instance* redefines the system, however infinitesimally, maintaining its present state or shifting its probabilities in one direction or the other (as each moment's weather at every point on the globe redefines the global climate). (1991:34)

Taking this view of system and instance, there are two parallel hierarchies of systems being built up. They are represented schematically in Figure 5.1. By saying something, a speaker (abbreviated S in the figure) produces an instance of linguistic form and an instance of verbal behavior (a speech act). As an instance of linguistic form, the utterance contributes to the system of how that speaker talks in situations like the current one. Going up the hierarchy, it also contributes to the system of how that speaker talks in general, how people like the speaker talk and to the language of the speaker's whole community. As an instance of verbal behavior, the utterance contributes to the definition of the speaker's behavior in situations like the current one. As a situated action in the world, the utterance is also part of the practice that is the speaker's identity and the fabric of the social category or categories that the speaker belongs to. Finally, the social landscape of the community is populated with such categories.[15]

language of S's community	↔	social structure of S's community
language of S's social category	↔	social content of S's social category
language of S in general	↔	S's identity
language of S in this kind of situation	↔	behavior of S in this kind of situation
linguistic form	↔	speech act

Figure 5.1: Parallel hierarchies of systems

Each level of linguistic patterning (the instance, the pattern of the speaker's speech in this type of situation, etc.) corresponds to a level of social patterning and is thus potentially meaningful. To illustrate this, let's take a particular example and examine it from the point of view of interpretation. Suppose that a middle manager named Kim gets upset and swears in a business meeting. The impact that swearing will have on Sandy, another employee present, will depend on Sandy's knowledge of how Kim usually behaves in business meetings. If Sandy has often heard Kim swear in that situation, then Sandy might know that Kim is only somewhat upset. If Kim almost

[15] The separation of the linguistic and social hierarchies in this discussion is somewhat artificial. For an utterance to count as part of the system of how people like the speaker talk, it has to be interpreted by that speaker and others against the existing ideologized social background and the speaker's position in it. Eckert (2000) suggests that iconic speakers and iconic speech acts can have more weight in the definition of such styles. Conversely, this implies that many speech acts or speakers can be deemed unimportant and therefore have little effect on the system.

never swears in that situation, to Sandy's knowledge, then Sandy might deduce that Kim is extremely upset. Further, business meetings are not the only situation in which Kim speaks to Sandy. If Sandy knows that Kim swears a lot, but not at business meetings, then the import of that instance of swearing (for Sandy) is different than if Sandy almost never hears Kim swear.

Moving up the hierarchy, Kim is busy constituting the social category middle manager for Sandy, Kim, and the others present. Thus Sandy will also interpret Kim's swearing in the context of the behavior of other middle managers and other employees of the company. Finally, the taboo nature of the swear words is constituted in part by the combined practice of all members of the speech community, and Kim, in swearing, is making use of (and perhaps loosening) this taboo. Note that it's not just the act of swearing that has all of these interpretations. Suppose Kim has a particular favorite swear word (and Sandy knows this), but chose a different one on this occasion. The 'bite' or shock-effect that Sandy perceives in Kim's swearing will depend on that choice, but also on the general level of taboo associated with the particular word at all of the successively larger parts of the speech community.[16]

The preceding paragraph talked about patterns of behavior, including patterns of behavior of entire communities. These patterns can only be meaningful to the extent that they are known by speakers (and therefore available to be interpreted and to use in the interpretation of new utterances). Of course, no speaker will have perfect knowledge of the speech patterns of the rest of his/her community, or indeed, even of his/her own speech patterns. What counts are the patterns that the speaker believes to exist. Such beliefs are based on the actual interactions that the speaker has had, but certainly extend beyond as the speaker extrapolates for his/her experience. As mentioned above, there is considerable potential for misunderstanding between speakers with different experience, who may or may not assume their position to be shared. Further, it is usually the prerogative of those in power to assume that their understanding of the patterns is indeed shared. With these caveats, I believe the

[16]For ease of exposition, this example has taken the notion of a 'business meeting' as a social situation to be invariant background, which of course it isn't. To the extent that Kim is a prominent or powerful individual, Kim's behavior also contributes to what a business meeting is in that company, and in general, for the people present.

model of parallel hierarchies of social structure and linguistic structure provides a basis for describing the relationship between instances of variants, patterns of variation, and social value.

Most variationist work focuses on the social value of the patterns, and not of the instance. In fact, Lavandera (1978) writes:

> [F]or cases of inherent variation it is reported that there are no speakers who never use a variant nor are there any who always use it. It is not therefore which form is chosen in any particular occurrence but the frequency with which one form is chosen over another alternative form which, when correlated with some other linguistic or extralinguistic element, takes on significance. (1978:174)

Lavandera contrasts this with cases of syntactic or lexical variation where each variant is marked as e.g., 'informal' or 'formal', and argues that the pattern of variants in formal or informal styles does not carry any meaning, but is rather an epiphenomenon derived from the fact that the forms are meaningful.

On the other hand, Coupland (in press) argues that in the study of style it's not aggregate patterns but individual situated acts that are most revealing.

> The basic problematic of style ... is inherently established at a local level which makes aggregation inappropriate, just as it would be inappropriate, say, to interpret musical performance through an aggregation of pitch or amplitude levels across performers and instances. (in press)

I would like to argue that both the pattern and the instance can be meaningful at the same time. I hope that it is clear from the business meeting example that patterns built up out of meaningful acts are not left uninterpreted but are called upon as the background against which new acts are interpreted. Further, mixing up formal and informal forms can create an intermediate level of formality—a pattern whose meaning is not identified with that of any of the forms that build it. Thus meaningful instances can combine to produce meaningful patterns. To extend Coupland's musical analogy, it is not possible to tell from only one note what key a piece of music is in. Likewise, in cases where the meaning was previously held to be primarily in the pattern, one can find evidence that instances are treated meaningfully as well. In the experiment reported in Chapter 4, the listeners attributed social value to individual occurrences of copula absence/presence.

Further, the experiment showed that the meaning of an instance of a variable can be modulated by the syntactic context of that variable. A speaker using copula presence in an unmarked environment (before NP) was judged (by the AAVE speakers) to be somewhat reliable/likeable/well educated, but not as much so as a speaker using copula presence in the marked environment (before V+*ing*). To use the terminology introduced in this section, using copula presence appears to count as evidence that a person is reliable/likeable/well educated for these speakers, or index these properties. But it counts as better or stronger evidence (or indexes the properties more emphatically) in precisely those contexts where it is unusual. Therefore, the experimental results indicate that non-categorical grammatical constraints, so far ignored in this discussion of social meaning, are not a separate issue from social meaning but rather intertwined with it.

To summarize, then, variation is (potentially) socially meaningful both at the level of the social import attached to individual variants, and to the social import of patterns of those variants. Further, non-categorical grammatical constraints can interact with the social meaning of a variant. In the remaining three sections of this chapter, I will discuss three existing approaches to modeling non-categorical grammatical constraints, asking how they can model these three aspects of the social meaning of variation.

5.2 Variable rules

The first attempt to model non-categorical constraints on variation was Weinreich et al.'s (1968) proposal of variable rules, extended and refined in Labov (1969), Labov (1972, Ch. 3), Cedergren and Sankoff (1974), and elsewhere. In Weinreich et al.'s original proposal, variable rules were rules in a community grammar, that is, descriptions of the linguistic behavior of a community. Labov (1969, 1972, Ch. 3) and Cedergren and Sankoff (1974), on the other hand, explicitly view variable rules as a part of competence grammar.

5.2.1 Variable rules as a part of linguistic competence

Section 3.3.1 of Chapter 3 gave two examples of variable rules from Labov (1972, Ch. 3). The second rule (deletion) is repeated here as Figure 5.2.

$$\begin{bmatrix}+\text{cons}\end{bmatrix} \longrightarrow \langle\phi\rangle \;/\; \left\langle\begin{matrix}{}^*\text{strid}\\ +\text{cons}\\ +\text{Pro}\end{matrix}\right\rangle \#\# \begin{bmatrix}\underline{\qquad}\\ -\text{nas}\\ +\text{cont}\end{bmatrix} \#\# \left\langle\begin{matrix}+\text{Vb}\\ +\text{Fut}\\ -\text{NP}\end{matrix}\right\rangle$$

Figure 5.2: Labov's variable rule of copula deletion

This rule is encoded in an extension of the formalism of Chomsky and Halle 1968 (SPE). It has an input (a consonant) that is rewritten as ϕ in a specific context. It differs from standard SPE rules in that it applies with some probability.[17] The specifications in angle brackets in the context indicate those features of the environment that favor the application of the rule. These features are given, from top to bottom, in order of how strongly they favor the rule. The '*' before the feature 'strid' indicates that the rule applies categorically after stridents. The specification +cons indicates that a preceding consonant (as opposed to a preceding vowel) strongly favors deletion. Below that, the specification +Pro indicates that a preceding pronoun (as opposed to a full NP) favors deletion somewhat. The specifications in the following environment are +Vb (verbs, including *gon*/*gonna*), +Fut (the additional effect of *gon*/*gonna* over other verbs), and −NP (predicate adjectives and locatives).[18] Rules like the one in Figure 5.2 are said to have a basic probability of applying (the 'input probability') which is either enhanced or muted by the specific features present in the environment.

Variable rules that are rules of community grammar include various social factors (i.e., a specification of the social groups a speaker might belong to) as additional probabilistic constraints on rule application. For example, the speaker's being a teenager rather than adult would increase the probability of application of the deletion rule in Figure 5.2. However, when such rules are viewed as rules of individual linguistic

[17] This is indicated in the notation by the angle brackets around the output of the rule.

[18] The possibility of following predicate nominals is not explicitly mentioned. The notational convention seems to be to leave out the least favorable environment, and to specify only those that have a relatively favoring effect.

competence, differences in rates of application across speakers are represented by giving different speakers different input probabilities for the rule.

The variable rules of contraction and deletion are part of a system of rules, some categorical and some variable, that Labov proposes as a partial model of AAVE phonology. In view of the interrelation between the two types of rules in his grammar, Labov writes:

> [T]he variable rules themselves require at so many points the recognition of grammatical categories, of distinctions between grammatical boundaries, and are so closely interwoven with basic categorical rules, that it is hard to see what would be gained by extracting a grain of performance from this complex system. It is evident that [both the categorical and the variable rules proposed] are a part of the speaker's knowledge of language . . . (1972:125)

Cedergren and Sankoff argue instead from the systematicity of the effects modeled by variable rules:

> [T]he notion of competence must be strengthened to include representations of systematic covariation between elements of language, even when this covariation cannot be described in categorical [. . .] terms. (1974:352)

Unlike much work in computational linguistics and psycholinguistics that makes use of probabilistic models and methods, Cedergren and Sankoff (1974), at least, appear to intend variable rules as a model of competence with a "random or nondeterministic component." While models such as those of Jurafsky (1996) or Bod (1998) incorporate probabilities to choose the most probable structure out of some candidate set, Cedergren and Sankoff see the probabilities as a random element in the system, positing a relationship between competence and performance that is "analogous to that between a probability distribution and a sample" (1974:353). Preston describes a similar model, using the analogy of a coin toss to describe the relationship between the underlying probabilities and the output:

> Variable items are unfair coins made so by appropriate weightings which derive from such characteristics of identity as age, sex, and social class[,] and by such features of the environment as formality, solidarity with other speakers, and power and status relations among interlocutors. (1989:20)

This point will be relevant to the discussion of the kinds of social meaning that variable rules can model in Section 5.2.2 below.

The proposed change to linguistic competence involves making it more concretely related to performance. Earlier conceptions of linguistic competence (as well as the current mainstream conception) view competence grammar as merely generating a set of well-formed sentence types (and their associated structures and semantic representations)—never mind how this knowledge actually interacts with the systems behind actual language use. The probabilistic view of competence, on the other hand, is couched as a system that gets used every time a sentence is produced. The underlying competence contains some weighted coins and the performance system must include a means of (metaphorically, of course) flipping those coins and incorporating the choices they make.

There have been two main challenges to this idea. The first is that probabilities are an improbable thing for people to know. This was put most emphatically by Bickerton:

> [I]n order that the average for his group should remain constant, the variation of the individual must be confined within a relatively narrow range. What keeps his percentages within those limits? And how can it keep within them unless something, somewhere is COUNTING ENVIRONMENTS and keeping a running score of percentages? . . . [S]peaker B must continually be saying to himself things like : 'Good Lord! A's percentage of contractions in the environment +V__ +__ NP has fallen to 77! I'll have to step up mine to – let's see: . . . what? About 86%!' (1971:460–461)

As has been pointed out many times, this argument is a non-starter. Bickerton has confused frequencies with probabilities. That is, what is in the grammar are probabilities, which don't need any history to operate properly. Coins don't remember how they've been tossed, and yet they manage to keep to about 50/50 heads/tails.

Still, one might wonder whether probabilities are something people can know, and how they might be learned. In fact, there is ample evidence that people do learn frequencies. For example, Grant et al. (1951) set up a task in which participants were asked to guess whether a light would flash in each trial. For different groups of participants, the light flashed none of the time, 25% of the time, 50% of the time, or

100% of the time. The participants ended up guessing at the same rate as the light actually did flash. For example, the participants in the 75% condition guessed that it would flash about 75% of the time, despite the fact that they would have been correct more often had they guessed it would flash every time. Haser and Chromiak (1977) provide evidence that 'frequency counting' as they term it is an automatic skill. In their study, participants were shown a list of words, including some repeats, and then asked to judge how many times each word appeared. They included 2nd, 4th and 6th grade students as well as college students in their study. They found that the 2nd graders did as well as the college students. Further, neither being told before seeing the list that they would be asked about frequencies nor practice and feedback on their performance improved the participants' accuracy. In all of these properties, frequency counting differs from other memory skills. Haser and Chromiak conclude from this that frequency counting is largely an automatic process, and suggest that it "is something that the organism engages in as an essential component of his processing of the world" (1977:182). Finally, learning mechanisms that convert frequencies in a training corpus to probabilities in a grammar are not hard to come by (e.g., Boersma and Hayes 1999). The question is not whether people can learn probabilities as part of their linguistic competence, but whether they do.

The other challenge to Labov and Cedergren and Sankoff's proposed revision to the notion of linguistic competence is that variable rules as they were originally formulated don't fit well with syntactic theories (as opposed to SPE phonology). In his essay "The quiet demise of variable rules" Fasold (1996)[19] points out that the formal mechanism of variable rules makes certain assumptions that are incompatible with syntax (including transformational syntax) and non-rule-based phonology:

> Unlike phonological variable rules, the syntactic transformations of [the Standard Theory and the Extended Standard Theory] did not convert a possible structure into another structure that was also possible in the language. Rather, they re-arranged an underlying structure so that it became more nearly sayable. ...
>
> There is a problem with trying to analyze [syntactic] variables as the consequence of a syntactic variable rule in the traditional sense of *syntactic*

[19] Originally published as Fasold 1991. Page numbers cited here refer to the 1996 publication.

> *rule*. There seems to be little reason to suppose that most of these variants are related by the application or nonapplication of an actual variable rule of syntax. (1996:88–89)

and later:

> Analysis of variation in syntactic structure, even when VARBRUL is used, is nonetheless not about "rules" at all, at least not in the ordinary sense. Rather, it seems to be about the social and discourse consequences of making certain choices within language ...
>
> In many cases, especially in syntax, but also in phonology, there is no motivation for tying variation to the rules of grammar—at least not to any rules of grammar that have been proposed so far. (1996:90–91)

As noted in §5.1.1 above, one of the difficulties in studying variation in syntax is the problem of analyzing the myriad choices that speakers make into discrete bundles of variants called variables. However, variation is tied to grammar (if not to the rules of grammar) by the very fact that grammatical factors have robust favoring or disfavoring effects on the distribution of variants in the variables that have been isolated and studied. Unless these effects can be given an explanation outside the grammar (see §5.4 below), some way must be found to formulate rules of grammar so that they are able to interface with this information.

In fact, that goal is not so elusive as it might seem. Probabilistic context free grammars attach a probability to each way of rewriting each non-terminal symbol (see, e.g., Charniak 1996). Other approaches extend this by allowing rewrite rules that cover more than one level in the tree (e.g., Bod 1998), by conditioning the probabilities for each way of rewriting the current node on how the tree had been built so far (e.g., Black et al. 1993), or by attaching probabilities to grammatical dependencies between words in the sentence (see e.g., Collins 1996). Probabilistic parsers of this sort can be used for disambiguation in automatic parsing, for speeding up parsing, and for predicting which string was actually said in voice recognition systems where the input is uncertain. (For an overview, see Manning and Schütze 1999, chapters 11 and 12.) Chapter 6 discusses how a sign based grammar can include information about the effect of the surrounding grammatical environment on variation.

To summarize so far, the variable rules proposal is essentially a proposal that the grammar include probabilistically applied rules constrained again probabilistically by features of the linguistic context. The next section considers how well this model can handle the various aspects of the social value of variation.

5.2.2 Variable rules and social value

The first aspect of the social value of variation is social meaning associated with styles, or patterns of frequencies of variants. This is the aspect that variable rules can handle most elegantly. Labov (1966) found a parallelism between social and stylistic variation in the variables he studied in New York City. When speakers from any socio-economic class are speaking carefully (as in a formal situation) they tend to adjust their rates of prestige variants towards that of speakers from higher socio-economic classes. Thus the stylistic range of any individual reproduces on a smaller scale the range of variation of the community. This is the 'attention paid to speech' model of variation, on which speakers have a natural vernacular that they use un-self-consciously and the ability to produce more prestigious styles when they pay close attention to how they're speaking. When this model is taken together with the idea of an input probability for variable rules, the obvious way to model stylistic variation with variable rules is to posit that speakers adjust their input probabilities according to the situation.

As discussed in Section 5.1.2 above, Labov's view of the relationship of style to social structure has been shown to be too simple in many ways. However, the variable rules model is not completely tied to the 'attention paid to speech' model of variation. First of all, speakers could be deliberately adjusting their input probabilities in either direction. Further, different variables could be associated with different dimensions of style, and their input frequencies adjusted independently of one another. Finally, speakers need not have one most natural input frequency, but could always be moving around within some range that they control.

	Interview III	Interview IV
Input probability	.49	.20
Following grammatical environment		
Gonna	[100%]	.79
V+*ing*	.59	.66
Locative	.87	.52
Adjective	.39	.29
Noun phrase	.24	.32
Miscellaneous	.34	.40

Table 5.1: Rickford and McNair-Knox's Varbrul input probabilities and factor weights for two interviews

As far as I know, this particular application of variable rules theory has not been explored extensively. However, it makes an interesting prediction: If stylistic variation is a matter of adjusting the input probability, then the constraining effects of features of the linguistic environment should remain the same across styles. Rickford and McNair-Knox 1994 is the only study I know of that reports the results of separate Varbrul[20] runs for different styles. The samples came from two different interviews with the same AAVE speaker, which represent different styles of speech primarily because in one case (Interview III) the interviewer was African-American and well known to the interviewee while in the other interviewer was white and not well known to the interviewee. The factor weights for the following grammatical environment are shown in Table 5.1, adapted from Rickford and McNair-Knox's (1994) Table 10.4, page 253. With the exception of locative in Interview III and adjective in interview IV, the relative orderings of the environments are the same across the two interviews, despite the overall higher rate of copula absence in Interview III. This is only one study of only one speaker, but it does suggest that the constraining effect of features of the linguistic environment is relatively stable across different styles or events.

The second aspect of the social meaning of variation is social meaning associated with individual instances. Here, variable rules do not fare so well. This was noted by

[20]As discussed briefly in Chapter 4, Varbrul is a computer program that implements the technique of stepwise multiple regression to isolate the effects of different factors on a linguistic variable. Varbrul is widely used without any presupposition of the theoretical notion of variable rules.

Dittmar (1996), who writes:

> Variable rules ... cannot account for different productive and interpretive communicative strategies because intention, underlying meaning, and pragmatic aims of communication are not considered in the analysis. (1996:134–135)

In other words, variable rules describe random, probabilistic behavior. If speakers produce one variant or the other probabilistically, then hearer must not attribute any significance to the choice between variants for any given instance.[21] Note that, if variable rules were just a model of performance—a statistical model of what people in a community do—that model wouldn't face this problem. But Labov and Cedergren and Sankoff insist that variable rules are intended as a model of probabilistic competence with a random component.

It might seem from the discussion of iconization above that hearers don't attribute any intentionality to speakers in their choice of linguistic variants. To return to the Japanese example, to use *wa* is to speak gently, and so women (who are supposed to be gentle) use *wa* as a matter of course. However, this is confusing ideologized social categories with actual people. Individual speakers are constantly navigating and negotiating their place in the social landscape, by actions such as choosing to speak like a woman (or like a certain kind of woman) or not in any given instance. Hearers must, at some level, recognize those choices on the part of their interlocutors.

Finally, for the third aspect of the social value of variation: the modulating effect of the linguistic environment on the social meaning of a variant. Variable rules cannot account for the effect shown in Chapter 4 because a theory that cannot attribute social meaning to individual instances has nothing to say about factors that affect the meaning of individual instances.

[21] This only applies to cases of inherent variation, that is, cases where speakers are known to actually use both variants, even in informal styles. In cases (should any exist) where each variant is strictly associated with one and only one style, one instance would be enough to determine the style in question. In that case, the social value of the style would be the social value of the instance.

5.2.3 Summary

The most important contribution of work done under the rubric of variable rules theory is the discovery of non-categorical grammatical constraints on variation. As a systematic aspect of language, non-categorical constraints require some explanation. The idea that non-categorical constraints could be a matter of probabilistic effects in a grammar is interesting and will be explored further in Chapter 6. However, the particular way in which the probabilities are incorporated in the model of competence/performance in variable rules theory appears to preclude an adequate treatment of the social value of variation. This, like the out-dated style of the phonological rules used, can safely be abandoned.

Before leaving this discussion of variable rules, I would like to consider one final criticism of them. Both Gazdar (1976) and Fasold (1996) point out that variable rules do not constitute explanations for the patterns they summarize. Further, as Fasold puts it, "nor does [a variable rule] do much to distinguish spurious patterns from significant ones" (p. 93).

Thus additional evidence that is independent of the production data is necessary in order to establish that information about frequency of co-occurrence is actually a part of speakers' internalized grammars. The experiment described in Chapter 4 was designed as a means of doing just this. As for the point that variable rules do not constitute explanations, no formal rules ever constitute explanations. A formal rule, as a part of a grammar, is a model, or description, of a language. In this sense, formal rules and formal grammars can account for patterns in language, but provide no information as to why they are that way. If we think of grammars as models of what people know (linguistic competence), then we can say that grammars go some way towards accounting for people's linguistic behavior. They do not constitute explanations of why a language is the way it is. Even in GB/Minimalism, where there is a emphasis on 'rich deductive structure' so that a small number of abstract properties of the grammar account for a wide range of linguist facts, these formal accounts only become explanations in connection with the theory of Universal Grammar *qua* Language Acquisition Device.

The next two subsections consider two different approaches to accounting for non-categorical constraints without adding them to the grammar. Section 5.4, in particular, addresses the possibility that non-categorical constraints can be explained by functional pressures on language use, and hence don't need to be represented in the grammar. But this assumes that grammars should only encode that which has no explanation. This assumption probably enjoys the wide acceptance that it does in part because of the relationship that does obtain between arbitrariness and grammar: if we speak according to arbitrary rules, we must know those rules. It does not follow from this that we don't also know or can't be shown to know (in the sense of represent as part of our grammars) other facts about language. This topic will be picked up again in Chapter 6. Variable rules (or some similar device) could be shown to correspond to something in actual speakers' knowledge of language without precluding the existence of some explanation for how the language that those speakers learned got to be that way.

5.3 Optimality Theory

Optimality Theory (OT) (Prince and Smolensky 1993) is a metatheory of grammar. The basic principles are simple and elegant. All grammars consist of the same violable constraints. Grammars differ in how they rank those constraints. The constraints chose grammatical forms from among a competing set of candidates. Between any two candidates, the winner is the one that does not violate (or violates fewer times) the highest ranked constraint on which they differ. For example, in the schematic tableau in Figure 5.3, candidate 1 is the winner.[22]

Within OT, a theory of grammar is a set of constraints, all of which are posited to be universal. The constraints define a 'factorial typology'—a set of rankings of the constraints which gives all possible languages based on those constraints.[23] The grammar of a particular language, in classical OT, is a complete ranking of all of the constraints. A complete ranking of constraints should choose only one optimal

[22] An '!' indicates the fatal violation, the one which causes an ungrammatical candidate to lose to the optimal candidate.

[23] And some definition of GEN, the function that gives the initial candidate set.

	CONS1	CONS2	CONS3
Cand 1	*		*
Cand 2	*	*!	

Figure 5.3: Example OT tableau

candidate out of any candidate set. There are two ways to get variation within a language in this system: either the variants correspond to different underlying forms so that the multiple optimal candidates do not ever compete, or the constraints are not able to distinguish the candidates. The first option is not very appealing for phonological variation, which is seen as differing realizations of the same underlying phonological form. The situation is not all that different for syntactic variables that do not correspond to semantic differences. The second option is also not appealing: in some cases of variation (i.e., changes in progress), one would expect to find speakers for whom one or another variant is completely unavailable. If some of the forms are to be ruled out for some speakers, then the constraints must be able to distinguish the candidates.

Crosslinguistic variation in OT is a matter of reranking the same universal constraints. Reynolds and Nagy (1994), Anttila (1997) and Boersma and Hayes (1999) have extended this idea to intraspeaker variation, by allowing speakers to have a range of possible rankings. The next three subsections describe their approaches.

5.3.1 Partially unranked constraints

Anttila (1997) provides perhaps the simplest extension to classical OT. In his framework, intraspeaker variation is a matter of speakers having partially unranked constraints in their grammars. Each full ranking of the grammar will still pick out only one candidate as optimal. The probability of occurrence of a given candidate is the proportion of full rankings that select it.

The particular phenomenon that Anttila investigates in this paper is the alternation between 'strong' and 'weak' variants of the genitive plural morpheme in Finnish.

Shorter words categorically take one or the other based on the weight of the final syllable of the stem.

(6) a. /puu/ 'tree' pu**i.den** strong
b. /kala/ 'fish' ka.lo.**jen** weak (Anttila 1997:37)

Words that are at least three syllables long and end in a light syllable, however, can take either the strong or the weak form. Further, this variation is subject to non-categorical constraints: (1) Stem-final high vowels favor the strong variant, while low vowels in this position favor the weak variant. (2) A heavy antepenultimate syllable (penult of the stem) favors the weak variant, while a light syllable in that position favors the strong variant. These results are based on native-speaker intuitions and a study of 1.3 million word corpus, and the results of both agree.

From a partial ranking of a set of plausibly universal constraints relating sonority, stress, and syllable weight, Anttila derives the categorical behavior of one and two syllable stems (and stems with heavy final syllables), and the possibility of variation in other stems. (This is possible because stress placement is more restricted in the shorter stems, there being less room to spread stressed syllables out.) The partial rankings are of a particular form: they divide the grammar into sets such that the sets are ranked with respect to each other, but the constraints within a set (when there are more than one) are mutually unranked.[24]

The probability of occurrence of each variant for the variable stems is determined by the number of mutually unranked constraints that each variant violates. (This is because the number of full rankings that select each candidate is directly related to the number of constraints violated.) In this sense, the non-categorical constraints are seen as reflecting the universal markedness of each form. In classical OT, the constraints are universal and all penalize some kind of markedness. Each language decides which dimensions of markedness it cares about more than others. This is the ranking of the constraints. In the case of unranked constraints, all of the constraints

[24]This is not quite true. Anttila's (1997) final grammar for Finnish includes a bottom set of constraints, some of which are mutually ranked because they are mutually ranked universally. It appears, however, that these constraints are ranked low enough as to play no role in the phenomena he considers.

count equally, and so there is a kind of additive markedness driving the distribution of the variants.

It is thus a prediction of this model that non-categorical constraints on variation reflect universal tendencies and thus should be the same whenever the 'same' variable arises (cf. the discussion in §5.4.2 below of the Kiparsky/Guy debate on this point).

One final property of this model that is worth mentioning is that the grain of the statistical predictions is limited by the number of unranked constraints participating in the system. If there are 6 constraints, then the smallest probability of occurrence possible for one variant (other than 0) is 1/6. Anttila notes this and suggests that a more complete grammar (one that considers more phenomena and thus has more constraints) should provide an even better fit for his data.

5.3.2 Floating constraints

Reynolds and Nagy (1994) also extend the OT treatment of crosslinguistic variation (different rankings of constraints) to sociolinguistic variation (in particular, to the variable deletion of segments and syllables after the main stress in a class of words in a Romance language called Faetar). Their approach differs from Anttila's in two important respects. First, at a formal level, instead of positing sets of mutually unranked constraints as Anttila does (7), they posit floating constraints that can move within some range (8).

(7)
$$\left\{\begin{array}{c}\textsc{cons1}\\ \textsc{cons2}\\ \textsc{cons3}\end{array}\right\} \gg \left\{\begin{array}{c}\textsc{cons4}\\ \textsc{cons5}\end{array}\right\} \gg \left\{\textsc{cons6}\right\}$$

(8)
$$\textsc{cons1} \gg \left\{\begin{array}{c}\ldots\;\ldots\;\ldots\;\textsc{cons2}\;\ldots\;\ldots\;\ldots\\ \textsc{cons3} \gg \textsc{cons4} \gg \textsc{cons5}\end{array}\right\} \gg \textsc{cons6}$$

If more than one constraint is allowed to float at a time, Reynolds and Nagy's system is a generalization of Anttila's, allowing for somewhat more fine-grained analyses.

The second respect in which Reynolds and Nagy's approach differs from Anttila's is in the data they attempt to account for:

> ... not only do the various rankings of the constraints account for each of the forms that surface and prohibit those forms that never surface, but the relative number of speakers who produce each surface form is correlated to the number of rankings which would produce each form. (1994:277)

That is, Reynolds and Nagy are not accounting for frequencies in some corpus of text, but for frequencies among a population of speakers. Their grammar is, in effect, a community grammar.[25]

5.3.3 Probabilistic OT

A third approach positing the possibility of multiple rankings within the same grammar is Boersma and Hayes's (1999) probabilistic ranking of constraints. On this approach, rather than strictly ranking the universal constraints, a grammar assigns each constraint to a point on a ranking scale. This point, called the ranking value, is the mean of a normal distribution of actual rankings of the constraint. That is, at evaluation time, each constraint is ranked somewhere near its ranking point, with the probability of any given point being determined by the normal distribution. All of the normal distributions are assumed to have the same standard deviation (width).

If two constraints A and B have exactly the same ranking point, then the chances of B outranking A in a given evaluation run are 1/2. If A is placed far above B on the ranking scale, then the chance of B outranking A in a given evaluation run can be made arbitrarily small.[26] Thus this system can simulate both the mutually unranked sets and the strict rankings of Anttila 1997. However, it can also allow for partial overlap of constraints, as fine-grained as necessary. While on Anttila's system, a distribution in which one variant occurred 1% of the time and the other 99% would require 100 constraints, Boersma and Hayes can model this with two constraints that overlap just a little bit, so that one outranks the other 99% of the time.[27]

[25] Nonetheless, they end up positing a slightly different partial ranking as the grammar of one particular minority dialect in the population they studied.

[26] According to Boersma and Hayes, if the ranking point of A is 5 standard deviations above the ranking point of B, the chances of B outranking A are 1 in 5,000

[27] Boersma and Hayes state that the 1:99 ratio could also be modeled in Anttila's system with only 5 constraints which interacted in such a way that only 1 of the 120 possible orders selected the uncommon candidate. This is inconsistent with Anttila's method of calculating the proportion of rankings that select each candidate. On Anttila's method, the proportion of rankings that select a candidate (and thus the predicted corpus frequency of a candidate) is directly related to the number of constraints that candidate violates (and that its competitors don't). If there are only five constraints, then the closet ratio to 1:99 possible is 1:4. This happens if one candidate violates only one constraint and the other violates only the other four. Perhaps this inconsistency has to do with the fact that Anttila's constraints don't interact in whatever way Boersma and Hayes have in mind.

This approach could be generalized to allow floating constraints, i.e., constraints that range over some series of other, strictly ranked constraints. In this case, some constraints would have larger ranges in which they vary (larger standard deviations). Boersma and Hayes opt for the more restrictive theory in which all constraints have the same range of variation, in the absence of evidence requiring floating constraints. (They note that they can account for Reynolds's (1994) and Reynolds and Nagy's (1994) data in their system, although not with the same constraints.)

5.3.4 Stylistic variation in OT

This section considers how the OT accounts could address the problem of stylistic variation—the fact that the distribution of variants can change in different styles, e.g., cases where the more prestigious variant is relatively more common in more formal or careful styles. One first pass at this problem, given the general framework of OT, would be to create different grammars (i.e., different (partial) rankings) for each style. Taking this as a reasonable strategy for now, how would it work in Anttila's (1997) system, the simplest of the three discussed here?

To be fair, Anttila (1997) states explicitly that he is trying to see how far a grammatical explanation can go towards accounting for all of the variability in his data:

> It is entirely possible that there exists variation which is not sensitive to styles, addressee, gender, age or socioeconomic class, but is completely grammar-driven. To what extent extragrammatical factors are needed in deriving accurate statistics remains an empirical question.

Therefore the exercise of extending Anttila's system to see how it would account for stylistic variation amounts to trying to make it do something it wasn't intended to do.

The final partial grammar for Finnish that Anttila settles on is as in Figure 5.4. For the class of trisyllabic stems with light final syllables, the constraints in the first two sets are unable to choose between the strong and weak forms of the genitive plural. The competition gets passed off to the third and fourth sets of constraints. Within the class of trisyllabic stems, Anttila distinguishes 6 subclasses (3 vowel heights x 2 weights of the penultimate syllable of the stem). In three of these subclasses, the constraints in the third set are decisive—that is, all of the constraints that apply pick the same candidate (strong or weak form). In the remaining three classes, either the constraints in the third set disagree on which form they select, or the constraints in the third set do not apply and the constraints in the fourth set disagree on which form they select.

SET 1	SET 2	SET 3	SET 4	SET 5
*X́.X́	*Ĺ	*H/I	*H/O	*H/A
	*H	*Í	*Ó	*Á
		*L.L	*L/A	*L/O ≫ *L/I
			*H.H	*A ≫ *O ≫ *I
			*H́	*L
			*X.X	

Key:

X.X	two successive syllables
X	an unstressed syllable
X́	a stressed syllable
H	a heavy syllable
L	a light syllable
A	a syllable with a low vowel
O	a syllable with a mid vowel
I	a syllable with a high vowel

Figure 5.4: Anttila's final partial grammar for Finnish

For ease of exposition I will abstract away from the content of the constraints, and rename the constraints in sets 3 and 4 as in Figure 5.5.

The first variable class of stems are those in which the penultimate syllable is light and the vowel in the final syllable of the stem is high. When combined with these

SET 3	SET 4
A	D
B	E
C	F
	G
	H
	I

Figure 5.5: Renamed constraints, sets 3 and 4

stems, the strong variant violates two constraints (A and B) and the weak variant violates only one (C) (cf. Table 5.2). There are six ways of ranking these three constraints, and in four of those rankings, one of A or B will be on top. This means that in 2/3 of the possible rankings of these constraints, the strong form will violate the highest-ranking constraint and the weak form won't. In the other two rankings, constraint C is on top, and strong form will be selected as optimal. Therefore, 2/3 of the possible rankings select the weak form for these stems. On Anttila's system, this predicts that 2/3 of the tokens of this kind of stem should appear with the weak form of the suffix.

form	constraints violated
strong	A, B
weak	C

Table 5.2: Constraints violated: light/high V stems

In the next set of stems, the penultimate syllable is heavy, and the vowel in the final syllable is low. These forms all satisfy the constraints in set 3 (since those constraints concern light penultimate syllables and/or high vowels), and the competition is passed off to set 4. The pattern of violation of the set 4 constraints for the strong and weak forms attached to these stem is as in Table 5.3: each form violates two of the four constraints in the set. By a similar calculation as above, this grammar predicts that each form should occur 50% of the time.

The final set of trisyllabic stems that give rise to variation have heavy penultimate

form	constraints violated
strong	G, H
weak	F, I

Table 5.3: Constraints violated: heavy/low V stems

syllables and a mid vowel in their final syllables. Again, the competition is passed off to set 4. The violations of the strong and weak forms for these stems are given in Table 5.4. Note that even within set 4 the constraints in play here are not the same as for the previous set of stems (Table 5.3). In particular, neither form in Table 5.4 violates constraint F and neither form in Table 5.3 violates constraints D or E. This is again due to the fact that the constraints concern properties not relevant to certain stems. For this class of stems, the strong form is predicted to appear only 20% of the time.

form	constraints violated
strong	D, E, G, H
weak	I

Table 5.4: Constraints violated: heavy/mid V stems

These predictions fit the corpus frequencies that Anttila found remarkably well (see Table 5.5). Anttila's (1997) statistics come from a study of a corpus of written standard Finnish. While this particular phenomenon may indeed be insensitive to stylistic factors, it is also possible that other registers of Finnish see a change in the distribution of the strong and weak forms of the genitive plural of the variable stems. Could this system model a hypothetical style in which the usage of the strong forms is increased across the board? Given the general framework, the difference between that system and the grammar in Figure 5.4 would have to be in the ranking of the constraints. The existing rankings are either presumed to be universal rankings or motivated by the categorical facts of Finnish morphophonology. Therefore, assuming that the hypothetical style we're trying to model is still consistent with UG and still maintains the categorical distinctions found in the style studied by Anttila, the only

Stem class	genitive pl form	predicted %	observed %
light/high V	strong	33	37.2
	weak	67	62.8
heavy/low V	strong	50	50.5
	weak	50	49.5
heavy/mid V	strong	20	18.2
	weak	80	81.8

Table 5.5: Predicted and observed frequencies of strong and weak forms

possible changes are further rankings of the as yet unranked constraints.

Note first that the constraints determining the frequency of occurrence of each form in the first stem class (light/high V) are distinct from those affecting the other two stem classes (compare Table 5.2 to Tables 5.3 and 5.4). This means that reranking the constraints could change the distribution of the forms in one stem class without affecting the other two. For example, adding the information that constraint B is dominated by constraint C would even up the distribution of strong and weak forms in the first class to 50/50, but change nothing else.

For the strong form to appear more than 20% of the time in the third class of stems (heavy/mid V) the single constraint violated by the weak form (I) would have to outrank some of the constraints violated by the strong form (D, E, G, H). Table 5.6 gives a sample of partial rankings and how they affect the distribution of strong and weak forms in the two stem classes affected by these constraints. The predicted distributions are deduced, as before, from the set of constraints each candidate violates. That information remains the same as in Tables 5.3 and 5.4. Demoting a constraint to a lower set keeps it from affecting the outcome for these candidates, so long as there remain relevant constraints in a higher set. The number of constraints in the higher set violated by each candidate determines the distributions.

Ranking 1 is the original ranking. Rankings 2 and 3 change the distribution of forms for one stem set but not the other. Ranking 4 increases the frequency of the strong forms in both stem class, but unevenly: from 20% to 33% in one and from 50% to 100% in the other. In general, it is not always possible to manipulate the frequency of the strong forms in each stem class independently, as the stem classes are sensitive

	ranking	strong/weak	
		heavy/mid V	heavy/low V
1	{ D, E, F, G, H, I }	20/80	50/50
2	{ F, G, H, I } $\gg$ { D, E }	33/67	50/50
3	{ D, E, G, H, I } $\gg$ F	20/80	33/67
4	{ D, E, F, I } $\gg$ { G, H }	33/67	100/0
5	{ D, E, F, G, I } $\gg$ H	25/75	67/33

Table 5.6: Partial rankings and predicted frequencies

to some of the same constraints. (In particular, given this system of constraints, it is not possible to increase the frequency of strong forms in the heavy/low V class without also affecting the heavy/mid V class.)

If the ranking in Figure 5.4 is indeed the grammar of one register of Finnish, and other styles/registers are a matter of ranking a few more constraints, then we predict a bunch of styles in which the frequency of the strong forms is changed for one stem class but not another, or changed in multiple stem classes but unevenly. Whether Finnish genitive plurals are used in stylistic variation at all, and, if they are, what patterns of frequencies occur are, of course, empirical questions.

Will the greater flexibility of Boersma and Hayes's (1999) system help? It turns out that the predictions regarding stylistic variation of Anttila's (1997) approach are partly due to the framework and partly due to the constraint set. Since the light/high V class does not violate any of the same constraints as the two heavy classes, this constraint inventory will always predict styles in which the proportion of strong forms is changed in the light/high V class and not the others, or vice versa.

To see what happens in a different kind of constraint set, I will posit some constraints for copula absence in AAVE.[28] In order to get the variation, there must be some constraint(s) pushing for copula presence and some pushing for copula absence.

[28] This is not necessarily the best OT analysis of copula absence in AAVE, and is at any rate simplified in its assumptions about the underlying form and the candidate set. Further, it is not designed to capture all of the categorical constraints on copula absence discussed in Chapter 3. It is only put forth for the sake of the argument.

Suppose that the relevant input forms are quasi-semantic forms giving the predicate-argument relations and the part of speech of the word that contributes the semantic predicate. Suppose further that present tense is not specified in the input (i.e., unmarked), while other tense/aspect forms are specified in the input. In this case, the constraint against copula presence (in present tense sentences) can be seen as a kind of a faithfulness constraint: putting *is* in the surface structure adds a feature (present tense) that wasn't there in the input.

(9) FILL(TENSE): Don't express more information (about tense) than is in the input.

We want the constraints against copula presence to treat the different following grammatical environments differently, so that they can be distinguished. A first pass would be the constraints in (10).

(10) *NP-PRED the predicate of a sentence must not be an NP.
*AP-PRED the predicate of a sentence must not be an AP.
*VPROG-PRED the predicate of a sentence must not be a progressive VP.

Using Anttila's (1997) system, if all four of these constraints are unranked, one would predict 50% copula absence regardless of environment, as shown in Table 5.7. For past tense forms, the input would include tense information (e.g., doctor_{NP}(Kim), past). Therefore, the presence of *was* in the output would incur no violation of FILL(TENSE), and the system would get categorical copula presence, if the input specifies past tense.[29]

input	copula absence violates	copula presence violates
doctor_{NP}(Kim)	*NP-PRED	FILL(TENSE)
happy_{AP}(Kim)	*AP-PRED	FILL(TENSE)
laughin_{Vprog}(Kim)	*VPROG-PRED	FILL(TENSE)

Table 5.7: Constraints violated: copula absence/presence

[29] Copulaless forms would remain underspecified for tense and thus could appear in both present and past tense contexts.

Using Boersma and Hayes's (1999) system, it would be possible to capture the observed difference in the frequency of copula absence in these environments. All one would need is a probabilistic ranking as in Figure 5.6.

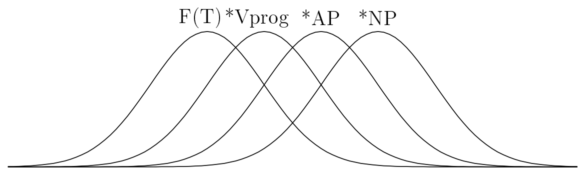


Figure 5.6: Constraints with overlapping distributions

The grammar represented in Figure 5.6 predicts copula absence less than 50% of the time in any environment, as Fill(Tense) is ranked lower on the scale than any of the other constraints. However, the distributions for Fill(Tense) overlaps with *Vprog-pred more than *NP-pred. This means that it will more often be the case that Fill(Tense) outranks *Vprog-pred than that it outranks *NP-pred. This, in turn, means that copula absence will be more common before V+*ing* than before NP. Note that the overlap of the three *X-pred constraints is not relevant, as those constraints won't ever apply to the same candidate set. Stylistic variation can be handled in this system simply by moving Fill(Tense) higher or lower on the scale.

There is the further issue of what these constraints predict in terms of possible languages. Among the possibilities are those given in Table 5.8.[30] This probably not the best possible set of predictions. System 4, in particular, looks like a very unlikely language. Suppose that the crosslinguistic patterns of copula absence follows

[30]The constraint *VProg-pred is somewhat problematic, as the copula is not always involved in the progressive crosslinguistically, even when there is a syntactically marked progressive and it is a compound tense. The constraint could be generalized to barring non-finite predicates. In any case, this issue is orthogonal to the immediate point of this section.

	Ranking	Property
1	Fill(Tense) $\gg \ldots$	categorical copula absence (present tense)
2	*NP-pred $\gg$ Fill(Tense) $\gg \ldots$	copula presence before NP, absence otherwise
3	*AP-pred $\gg$ Fill(Tense) $\gg \ldots$	copula presence before AP, absence otherwise
4	*Vprog-pred $\gg$ Fill(Tense) $\gg \ldots$	copula presence before V+ing, absence otherwise

Table 5.8: Predicted languages

the general ranking of constraints found in AAVE, as in (11).[31,32]

(11) ϕ V+ing $\succ$ ϕ AP $\succ$ ϕ NP

This is an implicational scale: a language with copula absence before NP will have copula absence in all other environments, etc. This state of affairs can be modeled by postulating a universal ranking of the constraints in (10), in particular, the ranking in (12).

(12) *NP-pred $\gg$ *AP-pred $\gg$ *Vprog-pred

The fourth constraint (Fill(Tense)) is ranked differently with respect to these other three in different languages, giving the range of systems described in (11). This is an example of Prince's (2000) 'element hierarchies', which contrast with 'inclusion hierarchies' as ways of modeling universal markedness. Prince is concerned with strictly ranked hierarchies, and doesn't consider the possibility of probabilistic rankings. It's possible that one might recast universal rankings such as (12) in terms of ranking points: it's okay for the constraints to be close together on the ranking scale, as long as their ranking points (midpoints) conform to the order given in (12).

For completeness, I will also discuss what happens if we posit a set of constraints that form an inclusion hierarchy. In an inclusion hierarchy, the more general constraints subsume the more specific ones (Prince 2000). In the case of (AAVE) copula absence, one might posit the following constraints:

[31]In fact, it probably doesn't. See §5.4.3 below.

[32]$\succ$ is OT for 'is better than' (of elements), and contrasts with $\gg$, which is OT for 'outranks' (of constraints).

(13) *NONFIN-PRED the predicate of a sentence must be finite.
*NONVERB-PRED the predicate of a sentence must be a verb.
*NP-PRED the predicate of a sentence must not be an NP.

*NONFIN-PRED is violated by any sentence in which the predicate is not a finite VP and *NONVERB-PRED by any sentence in which the predicate is not verbal. Any candidate that violates *NP-PRED also violates the other two, and any candidate that violates *NONVERB-PRED also violates *NONFIN-PRED. No implications go in the other direction. Therefore, this inclusion hierarchy describes the implicational scale given in (11) no matter how it is ranked. In particular, if we add FILL(TENSE) to the mix (still unranked), under Anttila's (1997) system, we get the predicted probabilities of copula absence shown in Table 5.9.

input	copula absence violates	copula presence violates	% absence
doctor$_{NP}$(Kim)	*NONFIN-PRED, *NONVERB-PRED *NP-PRED	FILL(TENSE)	25
happy$_{AP}$(Kim)	*NONFIN-PRED *NONVERB-PRED	FILL(TENSE)	33
laughin$_{Vprog}$(Kim)	*NONFIN-PRED	FILL(TENSE)	50

Table 5.9: Constraint violations and predicted frequencies

This looks more like the observed frequencies than did the predictions from the other constraints. Table 5.10 gives the predicted frequencies of copula absence by following grammatical environment for the possible partial rankings of these constraints on Anttila's (1997) system.[33]

Since the constraints *NONFIN-PRED, *NONVERB-PRED, and *NP-PRED form an inclusion hierarchy, they can only be crucially ranked with respect to each other

[33]The constraints are abbreviated in the table as follows:

F(T)	FILL(TENSE)
*NFP	*NONFIN-PRED
*NVP	*NONVERB-PRED
*NPP	*NP-PRED

	ranking	% copula absence		
		__ V+ing	__ Adj	__ NP
1	{F(T), *NFP, *NVP, *NPP}	50	33	25
2	F(T) ≫ {*NFP, *NVP, *NPP}	100	100	100
3	{*NFP, *NVP, *NPP} ≫ F(T)	0	0	0
4	*NPP ≫ {F(T), *NFP, *NVP}	50	33	0
5	*NVP ≫ {F(T), *NFP, *NPP}	50	0	0
6	*NFP ≫ {F(T), *NVP, *NPP}	0	0	0
7	{*NFP, *NVP, F(T)} ≫ *NPP	50	33	33
8	{*NFP, *NPP, F(T)} ≫ *NVP	50	50	33
9	{*NVP, *NPP, F(T)} ≫ *NFP	100	50	33
10	{*NFP, F(T)} ≫ {*NVP, *NPP}	50	50	50
11	{*NVP, F(T)} ≫ {*NFP, *NPP}	100	50	50
12	{*NPP, F(T)} ≫ {*NVP, *NFP}	100	100	50
13	{*NVP, *NFP} ≫ {*NPP, F(T)}	0	0	0
14	{*NFP, *NPP} ≫ {*NVP, F(T)}	0	0	0
15	{*NVP, *NPP} ≫ {*NFP, F(T)}	50	0	0

Table 5.10: Possible rerankings of inclusion hierarchy

if something else intervenes (Prince 2000).[34] As such, the only relevant rankings to consider are the rankings of Fill(Tense) with respect to the other constraints. Table 5.10 gives an exhaustive listing of such rankings. In the first one, none of the four constraints are ranked with respect to each other. This is the baseline. In the second and third rankings Fill(Tense) either dominates or is dominated by all of the constraints and copula absence or presence is categorical across all environments.

In the second set of rankings (4-6), one constraint is ranked above Fill(Tense) while the others can vary around it. This has the effect of decreasing copula absence, although not evenly across the board (except when they all go to 0). This kind of ranking can set copula absence to zero in one or more environments, according to the implicational scale encoded in the inclusion hierarchy.

The third set of rankings (7-9) are the mirror image of the second set. In this case, one constraint is ranked below Fill(Tense) and the rest. This has the effect of increasing copula absence, and ranking 9 comes the closest to an even, across-the-board increase in copula absence. Rankings 7 and 8 just change the values for one or two environments, respectively.

The final two sets of rankings (10-12 and 13-15) divide the constraints into two sets of two. Since Fill(Tense) is only ever unranked with respect to one other constraint in these rankings, there are only three possibilities: categorical copula presence, categorical copula absence, and 50/50 variation.

In general, these constraints have the following properties: Since they form an inclusion hierachy, under any ranking of the constraints, copula absence before V+ing will be at least as common as copula absence before Adj or NP, and copula absence before Adj will be at least as common as copula absence before NP. They define or predict three 'styles' with the strict ranking of environments shown in (11) (rankings 1, 4, and 9) and six (rankings 7, 8, 10-12 and 15) in which neither copula absence nor copula presence is categorical, but at least two environments are undifferentiated.

It seems that, as a theory of stylistic variation, this both predicts implausible styles and fails to predict enough styles. If speakers stylistically adjust their pattern

[34] If multiple constraints are relevant, they will always give the same result for any given candidate. For example, *nonfin-pred and *nonverb-pred will always treat candidates with adjectival predicates in the same way.

of variation according to the situation, one would expect more than three styles. Further, it seems unlikely that there should be a style of AAVE in which copula absence is possible and equally common across environments (ranking 10). However, Anttila's (1997) framework was not designed as a model of stylistic variation.

The addition of floating constraints (following Reynolds and Nagy 1994) doesn't change the predictions of the grammar significantly. The partial ranking in (14) is consistent with the four complete rankings in (15).

(14) $\left\{ \begin{array}{l} \ldots\ \ldots\ \ldots\ \text{F(T)}\ \ldots\ \ldots\ \ldots \\ \text{*NPP} \gg \text{*NVP} \gg \text{*NFP} \end{array} \right\}$

(15) a. *NPP ≫ *NVP ≫ *NFP ≫ F(T)
b. *NPP ≫ *NVP ≫ F(T) ≫ *NFP
c. *NPP ≫ F(T) ≫ *NVP ≫ *NFP
d. F(T) ≫ *NPP ≫ *NVP ≫ *NFP

Table 5.11 gives the grammaticality status assigned to copula absence in each of the following environments by each of the rankings in (15). Each ranking is assumed to be equally likely given the partial ranking in (14). The percentage of occurence of copula absence in each environment predicted by (14) is given in the last column of Table 5.11.

Environment	Ranking				% copula
	a	b	c	d	absence
__ V+ing	*	√	√	√	75
__ Adj	*	*	√	√	50
__ NP	*	*	*	√	25

Table 5.11: Predicted frequencies with floating constraints I

Even though the three ranked constraints are not in conflict, different partial rankings give different results. Table 5.12 gives all of the possible grammars with just one of these four constraints floating. Although there are more possible grammars and the details aren't quite the same, the range of possible values is similar to that found under Anttila's (1997) system. Again, because these constraints form an inclusion

Ranking		% Copula absence		
Floating	Ranked	__ V+ing	__ Adj	__ NP
*npp	f(t) ≫ *nfp ≫ *nvp	100	100	75
*npp	f(t) ≫ *nvp ≫ *nfp	100	100	75
*nvp	f(t) ≫ *nfp ≫ *npp	100	75	75
*nvp	f(t) ≫ *npp ≫ *nfp	100	75	75
*nvp	*npp ≫ f(t) ≫ *nfp	100	50	0
*npp	*nvp ≫ f(t) ≫ *nfp	100	0	0
*nfp	f(t) ≫ *nvp ≫ *npp	75	75	75
*nfp	f(t) ≫ *npp ≫ *nvp	75	75	75
f(t)	*npp ≫ *nvp ≫ *nfp	75	50	25
f(t)	*nvp ≫ *npp ≫ *nfp	75	25	25
f(t)	*npp ≫ *nfp ≫ *nvp	50	50	25
*nfp	*npp ≫ f(t) ≫ *nvp	50	50	0
f(t)	*nvp ≫ *nfp ≫ *npp	50	25	25
*nfp	*nvp ≫ f(t) ≫ *npp	50	0	0
f(t)	*nfp ≫ *nvp ≫ *npp	25	25	25
f(t)	*nfp ≫ *npp ≫ *nvp	25	25	25
*nfp	*nvp ≫ *npp ≫ f(t)	25	0	0
*nfp	*npp ≫ *nvp ≫ f(t)	25	0	0
*nvp	*nfp ≫ *npp ≫ f(t)	0	0	0
*nvp	*npp ≫ *nfp ≫ f(t)	0	0	0
*npp	*nfp ≫ *nvp ≫ f(t)	0	0	0
*npp	*nvp ≫ *nfp ≫ f(t)	0	0	0
*npp	*nfp ≫ f(t) ≫ *nvp	0	0	0
*nvp	*nfp ≫ f(t) ≫ *npp	0	0	0

Table 5.12: Predicted frequencies with floating contraints (all)

hierarchy, copula absence before V+ing is always at least as frequent as copula absence before an adjective, and copula absence before an adjective is always at least as frequent as copula absence before an NP. In addition to some rankings that give the strict ordering of environments, there are many others in which copula absence is equally frequent in two or even all three environments. Further, the stylistic range is again quantized. Since floating one constraint gives four possible rankings, copula absence (in any given environment) can only be predicted to happen at 0%, 25%, 50%, 75%, or 100%. Allowing up to three constraints to float would only allow for a few more levels of copula absence. There is no room for a style with slightly less copula absence, etc.

In Boersma and Hayes's (1999) system (which doesn't allow floating constraints), one could achieve a series of smoothly changing styles with grammars such as those in Figures 5.7 and 5.8. In these grammars, the constraints *nonfin-pred, *nonverb-pred and *NP-pred have completely overlapping distributions (represented as one distribution labeled 'others' in the figures). The ranking point for the constraint Fill(Tense) (f(t) in the figures) is moved depending on the style.

If Fill(Tense) has exactly the same ranking point as the other constraints, then the grammar is the same as ranking 1 of Table 5.10, and has the same predicted frequencies of copula absence: __ V+ing 50%, __ Adj 33%, and __ NP 25%. If the ranking point for Fill(Tense) is moved lower than the others, as in Figure 5.7, then the chance that Fill(Tense) is ranked below the other constraints increases, and the frequency of copula absence decreases evenly across the board. If it is moved far enough away so as to approximate a strict ranking (ranking 3 of Table 5.10), we get categorical copula presence.

Conversely, if the ranking point for Fill(Tense) is moved higher than the others, as in Figure 5.8, then the frequency of copula absence increases evenly across the board. If it is moved far enough away so as to approximate a strict ranking (ranking 2 of Table 5.10), we get categorical copula absence.

However, certain problems remain. The first is the range of grammars/styles predicted. These are partly defined by the constraints, so the grammars/styles predicted by these constraints when they are probabilistically ranked still include the odd styles

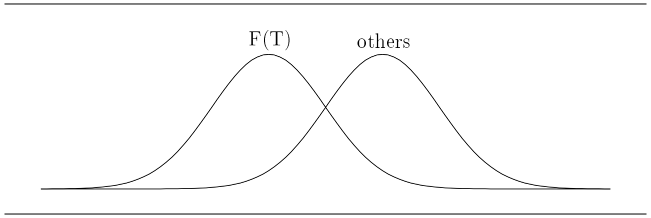


Figure 5.7: Style with less frequent copula absence

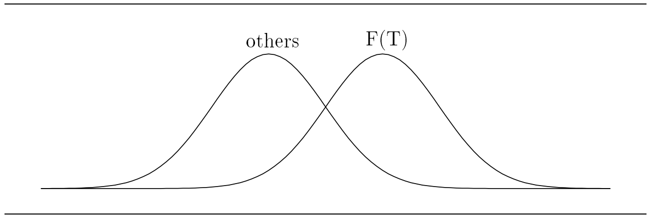


Figure 5.8: Style with more frequent copula absence

enumerated in Table 5.10. Second, perhaps the most interesting feature of Anttila's (1997) analysis of the Finnish data is that it derives the frequencies in the text from an underlying grammar that does not specify any frequencies. The analysis of stylistic variation in OT sketched here boils down to adjusting the position of one constraint along a continuous scale. It ends up looking fairly similar to a Variable Rules analysis: The ranking of environments inherent in the inclusion hierarchy corresponds to the non-categorical constraints on the Variable Rule, while stylistically adjusting the position of FILL(TENSE) corresponds to stylistically adjusting the input value.[35]

In fact, Boersma and Hayes's (1999) model is not only a model of grammar, but also the output of their Gradual Learning Algorithm. The Gradual Learning Algorithm is capable of abstracting a grammar (ranking constraints) from input including free variation, and of matching the frequency of each form in the input. This raises the question of whether the system sketched above for modeling stylistic variation is also learnable, and from what kind of input.

5.3.5 Variation and social value in OT

In this subsection I will argue that OT grammars such as those described above allow, to some extent, for an association of social value with a pattern of variation (a system), but not with individual forms used in individual utterances (an instance). The specific examples treated in this section will make reference to the grammars based on the inclusion hierarchy constraints (13), but all of the conclusions should apply equally to grammars based on the other set of constraints.

These OT models are based on the same kind of random application of probabilities as variable rules. For example, on Boersma and Hayes's system the exact ranking of constraints is determined for each evaluation by randomly selecting a point near each constraint's ranking point. Similarly, on Anttila's or Reynolds and Nagy's

[35]The correspondence would be less striking if *NFP, *NVP and *NPP didn't lead to grammars in other rankings in which copula absence is equally prevelant across all environments. Perhaps a better set of constraints could be found. On this note, one big contrast between the constraints I've been playing with for copula absence and Anttila's constraints is that Anttila's constraints were based on both plausible universals and categorical facts in the same language. Perhaps an OT analysis of categorical effects in AAVE syntax would provide some better constraints to play with.

system, the unranked constraints are randomly ranked for each evaluation run. If such a system is meant as a direct model of linguistic competence, it will have the same problem as the variable rules model did in terms of associating social value with particular instances: Since the choice of rankings and therefore of variants is random (according to specific probabilities), hearers can't attribute any intentionality to speaker's "choice" of variants. Without such intentionality, it's hard to see how there could be meaning.

Note also that it wouldn't do to associate social value with forms in the underlying representation. If any constraints sensitive to that social value are ranked highly enough, they will override the intricate interactions which provide the model of the non-categorical grammatical constraints. If there are no constraints sensitive to the social value information, or if such constraints are ranked low in the hierarchy, then the choice of forms will again be handled by the (randomly ranked) grammatical constraints. This means that even if a speaker picked an underlying representation with some social value, this choice could get overriden by the higher ranked constraints, producing some other variant. A hearer using the same system would have not way to know which variant the speaker actually intended.

Associating social value with a particular pattern of variation, on the other hand, is simply a matter of associating social value with a particular ranking of constraints. That is, speakers would have different grammars (rankings) for different styles. Now, it turns out that in some cases, associating social value with a grammar (i.e., a ranking) may be sufficient to associate social value with a particular form. For example, take a model where there are only a few discrete style-grammars, such as the rankings in Table 5.13, taken from Table 5.10. If these are the only rankings available, and a listener hears a speaker use copula absence before an NP, then the listener can deduce that the speaker must have been using ranking 1 and associate whatever social value goes with ranking 1 with the speaker's use of that one form. If the speaker uses copula absence before an adjective, then the listener can attribute a social value indeterminate between the values for rankings 1 and 4 with the speaker's use of that form, etc. But this won't apply if the range of style-grammars includes some that are not differentiated (like those in Table 5.13) in which forms they allow, but only

	ranking	% copula absence		
		__ V+ing	__ Adj	__ NP
1	{F(T), *NFP, *NVP, *NPP}	50	33	25
4	*NPP ≫ {F(T), *NFP, *NVP}	50	33	0
5	*NVP ≫ {F(T), *NFP, *NPP}	50	0	0
6	*NFP ≫ {F(T), *NVP, *NPP}	0	0	0

Table 5.13: A few discrete styles

in the probabilities with which they produce those forms.

What should a listener do when confronted with an overt copula? Since overt forms of the copula are consistent with all of the rankings in Table 5.13, nothing can be deduced about the grammar behind it. This means that speakers can't meaningfully use the overt copula at the instance level on this model. Such instances only contribute to the overall style.

There is at least one other way in which this system won't scale up: if two variables in a grammar are associated with independent social values. For example, suppose a language has variables associated with formality and different variables (subject to different OT constraints) associated with sounding confident. Certainly these dimensions are independent: one might want to sound more or less confident in both formal and informal speech. To model this under the system sketched here, one would need four rankings of constraints: the confident formal grammar, the confident informal grammar, the unconfident formal grammar, and the unconfident informal grammar. The number of style-specific grammars increases exponentially with the number of (differently socially valued) variables accounted for. Even if each variable only had two patterns of use (high and low use of one variant), two orthogonal variables require four grammars, three require eight, etc. The number of different levels of use within each variable gives the exponent of the exponential progression. In the limiting case of ranking on a continuum (as with the Boersma and Hayes (1999) model), the number of grammars is potentially infinite: for every different nuance of style someone could care to express, there is a different grammar.

The alternative is to associate the social value not with entire grammars (rankings) but with the relative rankings of small sets of constraints within the grammar. This isn't any different from associating social value with each variable directly.

5.3.6 Summary

In this section we have seen that the OT model can be generalized to account for variation and for non-categorical constraints on variation. However, the OT model is not well suited to handling socially meaningful variation. The most flexible version (Boersma and Hayes's (1999) probabilistic OT), can handle continuous stylistic variation. However, all OT accounts share with variable rules the view of variation as probabilistic, which entails that they do not allow any association of social value with individual instances of variants. Further, OT accounts still stipulate the ranking of environments: either by stipulating the universal hierarchies of constraints or by encoding the ranking of environments into an inclusion hierarchy. The most interesting feature of these OT models is that they connect the non-categorical constraints on variation to the analysis of categorical phenomena in the same language.

5.4 Functionalist accounts

The Variable Rules model holds that the constraining effect of different environments of a variable should be encoded directly in the grammar, as factors affecting the probability of application of a rule. The OT models encode categorical and non-categorical constraints in the same way: as violable, ranked constraints. Given the OT framework, allowing for variation (leaving constraints unranked) automatically makes some predictions about the frequency of each variant across environments. Functionalist accounts differ from both VR and OT in that the constraining effect of different environments is not stipulated in the grammar at all. The grammar provides free variation or optional rules, and functional effects in performance are behind the distribution of the variants across environments.[36] Kiparsky (1972) writes:

[36] This is not a pure functionalist model in that it assumes some formal grammar whose output is then subject to functionalist constraints.

> Labov (1969) has claimed that these frequencies are part of linguistic competence, and has developed a formalism for expressing them in grammatical rules. The alternative hypothesis would be that they are the result of general functional conditions impinging on speech performance. (1972:223)

As such, a functionalist account would be preferable to the other two in two respects: it leads to a simpler grammar and it actually provides an explanation for the ordering of environments. It is important to note that these two benefits of a functionalist account are independent. It is possible to give a functionalist explanation for constraints that are nonetheless part of the grammar. Indeed, in that same article, Kiparsky also states:

> Functional conditions, then, enter the linguistic system in a grammaticalized form. At that point they begin to interact and conflict not only with each other, but with formal generality in the usual generative phonological sense. (1972:224)

Further, his functionalist account of variation can be seen as the precursor to the OT accounts discussed above (Paul Kiparsky, p.c.). OT constraints are (ideally) functional, and the mechanism of OT is one way of formalizing and making precise their interaction. In particular, it provides a means of making explicit quantitative predictions.

Of course, OT is not the only possible model of the interaction between functional (and other) constraints. The remainder of this subsection will consider Kiparsky's (1972, 1988) functionalist proposal without the mechanism of OT. The comparison of this account with the variable rules model (and to some extent, with the OT model) can be seen as two questions: Are all of the non-categorical constraints on variation functional? Are any non-categorical constraints on variation encoded in the grammar? If some constraint on variation could be shown to be non-functional (or counter-functional), i.e., if the answer to the first question is 'no', then such a constraint would have to be encoded in the grammar. There would be no other basis for the observed effect in language production. However, even if all of the constraints on variation are functional, or can be understood as functional, this doesn't bar their being encoded in the grammar. It merely shifts the burden of proof to those who maintain that they are.

5.4.1 Social value on a functionalist account

One consequence of considering functonalist accounts as separate from their possible instantiation in an OT model is that they are not necessarily subject to the same shortcomings with respect to modeling the social value of variation.

First, there is nothing blocking the association of social value with individual forms in a functionalist account, since it need not assume that the choice between forms is (entirely) probabilistic. Since styles don't correspond to particular grammars, they can be as fine-grained as necessary, simply built up out of a speaker's linguistic choices over some stretch of time.

Further, if listeners can assume that speakers used one form or the other on purpose, it is possible to account for an interaction between non-categorical constraints and social value. If a speaker uses a variant in a disfavoring environment, then a hearer can deduce, via Gricean reasoning, that the speaker probably really wanted to convey the social value associated with that variant. Consider final t/d deletion, the example that Kiparsky focussed on. Deletion of t/d is inhibited when the /t/ or /d/ is the sole reflex of a morpheme, as in regular past tense. Assuming that this variable is associated with some social value, deleting the final /t/ or /d/ of a regular past tense verb must mean that conveying the social value was important enough to risk the possibility of the listener mistaking the past tense verb for a present tense form. Similarly, t/d deletion is sensitive to the sonority of the following segment (i.e., the first segment of the following word). Retaining t/d even when the following segment has low sonority would mean that the speaker went to extra articulatory trouble to be sure to pronounce the final /t/ or /d/. Again, a listener could interpret this to mean that the speaker especially cared to convey the social value associated with retaining t/d.

Note that this explanation for the effect of context on social value is dependent on the listener having and exploiting either knowledge of the functional constraint

in question, or knowledge of distribution of variants across contexts. That is, while the production data alone may be consistent with a functionalist explanation that requires no direct representation of non-categorical constraints, the perception data show that speakers must know the constraints, in one form or another.

5.4.2 Empirical tests

The importance of Kiparsky's (1972, 1988) work is that he goes beyond an appeal for functional explanation and presents a concrete proposal. He proposes that variation is subject to a functional constraint called the distinctness condition: the application of phonological rules is inhibited in environments where their application would erase morphological distinctions. This constraint predicts the basic facts of t/d deletion: it is most common in monomorphemes like *most/mos*, less common in irregular pasts likes *kept/kep* and still less common in regular pasts like *steeped/steep*. (Some frequencies are given in Table 5.14, from Guy 1996:227.) Guy points out that while the distinctness condition directly predicts the contrast between monomorphemes and regular past forms, it can only predict the intermediate status of irregular past forms if it is interpreted as a scalar constraint: "the higher the functional load on a class of targets, the more resistance they exhibit to a process that would eliminate them." (1996:228)

	N	% Deleted	Factor weight
Monomorphemes (e.g., *mist, bold*)	658	38.1	.64
Irregular Past (e.g., *lost, told*)	56	33.9	.55
Regular Past (e.g., *missed, tolled*)	181	16.0	.23

Table 5.14: Guy's English coronal stop deletion by morphological class

Because Kiparsky attempts an explain the facts, and not just account for them, his proposal is subject to a stringent test. If a functionalist account is to be a true explanation of a phenomenon, then the same functionalist pressures must be active in every language. This is where OT accounts using functionalist constraints differ from the performance-based functionalist account considered here. In OT, all constraints are active in all languages, but in different languages different constraints can be more important. This is possible because the constraints are in the grammar, and their ranking is something that speakers learn.

Guy (1980) took up Kiparsky's (1972) challenge to a variable rule account of t/d deletion.[37] Guy argued that if the same grammatical factor could have differing effects on a variable in different varieties, this would show that speakers must learn such non-categorical constraints and that they are therefore a part of the grammar. Indeed, Guy found just such a situation by comparing the constraints on t/d deletion among white speakers in New York and Philadelphia. As shown in Table 5.15,[38] while the effect of a following consonant or vowel was similar in both varieties, the effect of a following pause was radically different.

Environment	__ #C	__ #V	__ ##
New Yorkers	100%	56%	83%
Philadelphians	100%	38%	12%

Table 5.15: Coronal stop deletion by following segment

Kiparsky (1988) replied by pointing out that the difference in the treatment of a following pause in the two varieties might follow from some other difference. He illustrates this with a hypothetical example: Suppose that t/d deletion were to only apply to unreleased stops, and that word-final stops were obligatorily unreleased in word final position before a pause (and elsewhere) for New Yorkers while they were only optionally unreleased for Philadelphians before a pause.[39] Since this difference affects the input to the t/d deletion rule, it would account for the different ordering of the following environments for that rule. With this argument, Kiparsky showed that, without controlling for other, potentially relevant differences in the dialects, Guy

[37] In particular, to Labov et al.'s (1968) account of t/d deletion.
[38] From Kiparsky 1988:387.
[39] Kiparsky seems to assume that the hypothetical difference in the environments for unreleased stops is somehow different from Guy's difference in the environments for t/d deletion. For the unreleased stops, it's a difference between a rule applying categorically (New York) and applying optionally (Philadelphia) in the environment in question, as opposed to applying optionally, but at different rates in the different varieties. However, categorical application can be seen as the limiting case of frequent, optional application.

(1980) had not in fact provided a counterexample to the functionalist explanation of this phenomenon.

Guy (1996) presents two more arguments against Kiparsky's (1972) functional proposal. In order to discuss these, it is necessary to first introduce one more aspect of the original proposal. Based on a survey of historical changes in which phonology did or did not affect morphological categories, Kiparsky (1972) develops the following typology of morphological categories:

(16)

Weak	Strong
verb agreement	verb agreement
(languages without pro-drop)	(languages with pro-drop)
case	number
	tense
	gender

The strong categories are those that bear more functional load, the weak categories those that bear less. The systems of grammatical case that Kiparsky studied and verb agreement in languages without pro-drop carry only recoverable information.[40] Gender is the odd one out here, behaving diachronically as a strong category although the grammatical gender of a noun should always be recoverable from the lexical entry for the noun.[41]

Guy's (1996) first argument has to do with past participles in English. For regular verbs, the past participle has exactly the same form as the past tense. In the past participle, however, the suffix bears only redundant information, according to Guy.[42] Nonetheless, past participles are deleted infrequently, just like regular pasts, as shown in Table 5.16, from Guy 1996:231.[43]

Guy argues that the distinctness condition should not prevent t/d deletion in past participles: if the distinctness condition can take into consideration the larger context,

[40] One would not expect case to be a weak category in the Australian languages studied by Nordlinger (1998).
[41] The exception is homophones with different gender assignments, such as *die Stadt* 'the city' and *der Staat* 'the state' in German (Thomas Wasow, p.c.).
[42] Actually, given that the participle can appear without *have*, perfective aspect is better analyzed as being expressed by the participle. Guy does not differentiate the data he presents according to the syntactic context of the participle.
[43] The results for the three categories also shown in Table 5.14 are not exactly the same here, as these results come from a different corpus.

	N	% Deleted	Factor weight
Monomorphemes	739	38.6	.64
Irregular Past	74	35.1	.60
Regular Past (e.g., *talked*)	157	19.1	.41
Past Participles (e.g., *have talked*)	74	17.6	.35

Table 5.16: Guy's Cornonal stop deletion in 4 morphological classes

the information is redundant, and even if it can't, "'participle' is not one of the strong categories that resist erosion." (1996:230) This last argument is less than convincing, however, since the participle is part of the tense/aspect system of English, so it might be expected to fall under the 'tense' category in Kiparsky's typology. A better interpretation might be that the distinctness condition predicts deletion in participles to be intermediate, as it is for irregular past tense verbs. In both cases, the final t/d does (arguably) represent a morpheme, although one that bears only redundant information. One might object that there is a difference in that for the irregular pasts the final t/d is redundant within the word, while in the participles it's only redundant within the sentence. However, note that Kiparsky's basis for classifying case as a weak morphological category was that it supplies information already contributed by something else in the sentence, such as word order or a preposition. Under this interpretation, the participles do constitute a counterexample to Kiparsky's proposal.

Guy's (1996) second argument has to do with denasalization in Popular Brazilian Portuguese (PBP). The morphological category threatened in this case is number agreement within 3rd person verb forms. For regular verbs, the only difference is nasalization of the final vowel:[44]

(17)

3rd sg	3rd pl	gloss
fala	falam	'speak'
sobe	sobem	'go up'

Guy's results for denasalization by morphological class are given in Table 5.17.

At first glance, this looks wildly counterfunctional: regular plural verbs are denasalized more than monomorphemes. However, the picture is somewhat complicated

[44] Rendered orthographically with a ˜ over the vowel, or with a following *m*.

	% Denasalized	N
Noninflectional nasality (e.g., /õmĩ/ 'man')	66	1258
'Irregular' plural verbs (redundant marker) (e.g., /fizerũ/, cf. sg /fez/)	67	1446
Regular plural verbs (unique marker) (e.g., /sobĩ/, cf. sg /sobi/)	85	3783

Table 5.17: Guy's Denasalization in PBP

by the fact that, in addition to the phonological rule in question, the syntax of PBP makes subject-verb agreement optional in the first place. Guy estimates the rate of underlying agreement by considering forms that are not subject to denasalization (forms with final stress, which blocks the rule), and finds that syntactically, verbs agree 63% of the time. Assuming that syntactic agreement occurs at the same rate in all verbs, the rate of denasalization given for regular plural verbs in Table 5.17 would only fall to 76%, still higher than the rate for monomorphemes.

This argument is based on the premise that 'number' is a strong category. However, the data Guy actually examines fall under the 'agreement' category, which is strong or weak depending on whether the language has pro-drop. Guy does not indicate if PBP has pro-drop or not. I interviewed three speakers from Rio, and found that their variety at least does have pro-drop:

(18) a. Já comi
Already eat-1sg.past
'I already ate.'

b. Já comeu
Already eat-3/2sg.past
'She/he/you already ate.'

Interestingly, there is some syncretism in the person/number inflection, such that (18b) is actually ambiguous between third and second person. This suggests that pro-drop in Brazilian Portuguese might bear some similarity to Japanese/Chinese pro-drop, which is not supported by any person/number inflections on the verb. If this is true, then this casts some doubt on the treatment of agreement as a strong category in Brazilian Portuguese.

However, even a knock-down argument against a given functional explanation of some non-categorical constraint would only be an argument against that one functional explanation. Likewise, a solid demonstration of different orderings of environments in different varieties would only show that that variable was not constrained functionally. These would be interesting results, as they would show that grammars can include this kind of information. However, just as functional explanations don't rule out grammaticization, the fact that some non-categorical constraints are grammaticized wouldn't rule out the possibility that others are actually purely functional. So what is known that might bear on the issue of whether the effect of the following grammatical environment on AAVE copula absence is functional? Two kinds of evidence are available in the literature: studies of variable copula absence in AAVE and other Englishes/creole varieties, and information about the copula in other languages. Both kinds of data are relevant, as any functional constraint posited to account for the data in AAVE should be potentially apparent cross-linguistically.

5.4.3 A functional explanation for copula absence?

Rickford (1998) provides a summary of copula absence by following grammatical environment in a number of varieties bearing different relationships to AAVE. His table 6.16 (1998:190) is reproduced here as Table 5.18.[45] The historical attestations are studies of "ex-slave recordings," early recordings of African Americans born in the mid to late 1800s. The diaspora recordings are modern recordings of speakers descended from African Americans who left the US to settle in Nova Scotia, Liberia, and the Dominican Republic. The varieties spoken by these speakers are thought to be sister varieties to modern day AAVE, and have been relatively isolated so that they bear little influence from other languages. The creole varieties studied are all English-based creoles, mostly with African substrates. The white American English varieties are all from the South, where one would expect mutual influence with AAVE.

[45] In this table, percents are marked with %. All other figures are Varbrul weights.

	NP	Adj	Loc	V+ing	gon
Historical attestations					
Ex-slaves (Bailey 1987)	12%	29%	15%	71%	100%
Ex-slaves (Poplack and Tagliamonte 1991)	.39	.27	.67	.72	.78
Diaspora recordings					
Samaná (Poplack and Sankoff 1987)	.41	.19	.23	.46	.59
Samaná (Hannah 1996)	.12	.44	.42	.89	.93
ANSE[a] (Poplack and Tagliamonte 1991)	.31	.46	.49	.69	.73
LSE[b] (Singler 1991) Carolina	43%	93%	100%	97%	100%
LSE[b] (Singler 1991) Albert	32%	65%	94%	85%	100%
LSE[b] (Singler 1991) Slim	36%	43%	91%	79%	100%
Creole varieties					
Hawaiian Creole (Day 1973)	63%	72%	62%	94%	No data
JC[c] 1960 (Rickford 1996a)	28%	81%	18%	86%	100%
JC[c] 1991 (Rickford 1996b)	4%	59%	28%	58%	93%
Bajan 1980s (Rickford and Blake 1990)	.08	.42	.54	.65	.77
Bajan 1991 (Rickford 1992)	.07	.71	.52	.89	1.00
NSLE[d] (Singler 1991) basilect	20%	92%	23%		
NSLE[d] (Singler 1991) mesolect	93%	100%	100%		
NSLE[d] (Singler 1991) acrolect	5%	13%	0%		
Trinidadian groups (Winford 1992)	1%	79%	90%	94%	97%
Trinidadian individuals (Winford 1992)	1%	30%	53%	70%	50%
White American English					
White Mississippi *are* (Wolfram 1974)	31%	49%(Adj/Loc)		66%	86%
White Mississippi *is* (Wolfram 1974)	8%	16%(Adj/Loc)		18%	18%
White East-Texas (Bailey and Maynor 1985)	2%	10%	8%	34%	54%
African-American Vernacular English					
is, NYC Thunderbirds (Labov 1969)	.2	.48	.36	.66	.88
is, NYC Jets (Labov 1969)	.32	.36	.52	.74	.93
is, NYC Cobras (Baugh 1979)	.14	.72	.31	.59	.78
is+are, Detroit WC (Wolfram 1969)	37%	47%	44%	50%	79%
is, Los Angeles (Baugh 1979)	.32	.56	.29	.66	.69
are, Los Angeles (Baugh 1979)	.25	.35	.69	.62	.64
is+are, Texas kids (Bailey and Maynor 1987)	.12	.25	.19	.41	.89
is+are, Texas adults (Bailey and Maynor 1987)	.09	.14	.15	.73	.68
is+are East Palo Alto (Rickford et al. 1991)	.29	.47	.42	.66	.77

[a] African Nova Scotian English, [b] Liberian Settler English, [c] Jamaican Creole, [d] Non-Settler Liberian English

Table 5.18: Rickford's Summary of copula absence rates

The last part of the table gives the results of several different studies of AAVE around the US. Thus the varieties represented in Table 5.18 are not unrelated to each other.[46] Their very relatedness should work with any functional pressure towards a uniform ordering of the following grammatical environment.

It is clear from this table that there is general agreement as to the ordering of environments. However, the agreement is not perfect. Most prominently, there is a striking lack of agreement on the ordering of the Adj and Loc environments. Among the datasets listed in this table, 15 have more copula absence before Adj than before Loc, while 12 have more copula absence before Loc than before Adj. One might attribute this to some functional constraint (cf. *NONVERB-PRED in the OT analysis above) which doesn't distinguish between these two. The variation in their ranking across the datasets would be the result of statistical noise. Such an explanation would be stretched, however to account for the large differences between Adj and Loc for Liberian Settler English (Albert and Slim) (65%/94% and 43%/91%), Jamaican Creole (1960) (81%/18%), Non-Settler Liberian English basilect (92%/23%), and the NYC Cobras *is* (.72/.31).

Other outliers include the high rate of copula absence before NP for ex-slaves in Poplack and Tagliamonte's (1991) study, Samaná speakers in Poplack and Sankoff's (1987) study, and for *is* for Los Angeles speakers in Baugh's (1979) study; the high rate of *are* absence before adjectives in Baugh's (1979) study; and the relatively low rate of copula absence before V+ing for Liberian Settler English speakers Albert and Jim (Singler 1991).

To summarize, Table 5.18 shows both striking similarity and non-trivial differences across datasets, and both need to be explained.[47] There are two general possibilities involving functionalist explanations: (1) A functional constraint is at work in all of the synchronic grammars, but its effects are distorted in some varieties by some other difference between the varieties (cf. the discussion of t/d deletion and unreleased

[46] Rickford is researching whether AAVE has creole origins, using the creole varieties as a point of comparison and the historical attestations and diaspora recordings as potential sources of information about earlier stages of AAVE.

[47] Although some of the difference between the data sets could be due to differences in e.g., coding across the different studies.

stops above). (2) A functional constraint was involved in the diachronic development of these systems, but it got grammaticized, opening up the possibility for arbitrary, learned differences between languages.

One possible kind of functional constraint (and the one tacit in the OT constraints given in §5.3.4) has to do with the canonical part of speech of predicates. Since finite verbs are the canonical predicates, one might expect that the further a category is from finite verb, the more likely the language would 'help' it become a predicate—with a copula of some sort. The underlying reason for this could have to do with canonical predicates carrying tense information, and tense information having more affinity with verbs semantically. If there is a functional constraint along these lines, one would expect to find it reflected cross-linguistically. To put it differently, functional constraints, not being learned, should be universal. Any universal constraint should be able to leave its mark on language evolution, because language change progresses through variation. Functional constraints on variation should therefore favor particular systems in the long run, and the effect of a functional constraint should be apparent in typological study. According to the proposed constraint, languages should only allow NP predicates without copular support if they also allow adjective and locative predicates without copular support, and those only if they also allow non-finite verbs to head sentences.[48] Indeed, any functional constraint capable of producing the distribution of coupla absence in AAVE should produce the same cross-linguistic typological pattern.

To test this prediction, I studied the reports on 18 languages prepared by John McWhorter under the auspices of the copula project at Stanford University.[49] These reports provide an overview of copular constructions (what Hengeveld (1992) calls 'non-verbal predications') in: Bambara, Bengali, Ewe, Finnish, Gã, Haitian Creole, Hindi, Hungarian, Igbo, Indonesian, Jacaltec, Nahuatl, Nama, Sranan, Swahili, Vietnamese, Yagaria, and Yoruba. The information was collected from grammars

[48] Of course, it's always possible that other factors have intervened in the development of the languages that don't fit the pattern. But if we can't observe its effect cross-linguistically, how do we know that a proposed constraint is really an extragrammatical functional factor and not grammaticized in the language in question?

[49] "Copula Contraction and Absence in Vernacular Black English and Other Varieties", NSF grant BNS-8913104, PI John Rickford.

(usually multiple sources per language) and some consultation with native speakers. In addition, Hengeveld (1992) presents a typological study of non-verbal predication. Among the languages in his sample, the following allow copulaless clauses with NP, AP or locative predicates:[50] Babungo, Chukchee, Gilyak, Guaraní, Hixkaryana, Ibambura Quechua, Jamaican Creole, Ngalakan, Ngiyambaa, Tagalog, Tamil, Thai, and Yessan-Mayo.

The first thing that is clear from this data is that the set of environments for *be* in various Englishes is not necessarily treated uniformly in other languages. There is a tendency for what McWhorter calls the "equational" and "attributional" (roughly NP and AP) sentences to involve a common form of the copula, while locatives use a different formative. Progressive is, of course, not always a compound tense, and even when it is, the auxiliary involved may or may not be the same formative used in the other copula environments. Nonetheless, it is still possible to look at which predicates require the support of some formative across languages.

Not all of the languages considered have (or were reported to have) compound tenses. When a language did have compound tenses (some including a progressive), the auxiliary in the compound tenses was always obligatorily overt. Of course, a language without compound tenses might be simply allowing non-finite verbs to head sentences. To determine whether this is the case would require close study of each language to determine the tests for finiteness.

Turning now to NP, Adj and Loc, we find the following patterns. Sranan (a creole spoken in Suriname) has an obligatory copula with NP, and an optional copula with AP and Loc in the present tense. This pattern supports the proposed functional universal. Similarly, Jamaican Creole and Babungo (a Niger-Congo language spoken in Cameroon) allow locative predicates without the copula, but not APs or NPs. These are the only languages in the data set that support the functional universal (although many others don't contradict it, see below). Note that Sranan and Jamaican Creole share with AAVE many similarities in their development as creoles. Further, Babungo comes from the region that supplied the substrate languages for these creoles.

[50] This list does not include languages that only allow Hengeveld's (1992) 'zero-1' copulaless clauses, those where the predicate bears verblike morphology.

Haitian Creole and Tagalog (Austronesian, Philippines[51]) simply have no copula across the board. Ngiyambaa (Pama-nyungan, SE Australia), Ibambura Quechua (Andean, Ecuador), and !Xu (Khoisan, Namibia) have optional copula absence across the board. As such, these languages would be compatible with any implicational scale. Also in this category are Tamil (Elamo-Dravidian, South India and Sri Lanka), Yessan-Mayo (Sepik-Ramu, North New Guinea), Hixkrayana (Ge-Pano-Carib, North Brazil), and West Greenlandic (Eskimo-Aleut), which allow copula absence with locatives. For various reasons, the Adj and NP environments are not relevant in these languages.

Bambara (Niger-Congo, Mali) has an optional copula in "equational" (NP-pred) sentences and an obligatory copula (of a different form) elsewhere. McWhorter gives these examples, from Brauner 1974:[52]

(19) Bambara:

a. O ye lakòli ye.
that COP school COP
'That is the school.'

b. N'tògò ye Camara.
my-name COP Camara
'My name is Camara.'

c. N'tògò Fatumata Koné.
my-name Fatumata Koné
'My name is Fatumata Kone.'

d. Jiri ka kóró.
tree COP tall
'The tree is tall.'

[51] Language family and locations for language are as in Hengeveld 1992 or *Ethnologue*, Barbara F. Grimes, ed., web edition, http://www.sil.org/ethnologue/

[52] Rickford (1998) notes that copula absence examples in sentences giving names (e.g., (19b)) is less than clear-cut, as many languages have verbs meaning 'be named' which may have the same form as the noun 'name'. Nonetheless, it is interesting that (19b-c) show that the copula is variable in this environment in Bambara.

e. Mobili bè yan.
Car COP here
'The car is here.'

Similarly, in Jacaltec (Mayan, Guatemala and Mexico), sentences with NP predicates never take a copula, although those with AP predicates do in some cases. McWhorter gives the examples in (20), from Craig 1977.[53]

(20) Jacaltec:
a. Sonlom naj.
marimba player he
'He is/was a marimba player.
b. C'ul ix.
good she
'She is good.'
c. Itz'at ye ix.
smart COP she
'She is smart.'
d. Ay-ah-toy naj swi' te' nah.
COP-up-away he its-head the house
'He is on the roof of the house.'

Abstracting away enough to compare, both Bambara and Jacaltec have a distribution of copular forms that runs exactly counter to that found in AAVE.

A more common pattern, also at odds with the AAVE pattern, is for locatives but not NP or adjectival predicates to appear with copula support, either optionally or obligatorily (usually only in the present tense). Here I am counting as optional copula absence both free variation and variation associated with semantic differences such as temporary/permanent state. This pattern shows up in Bengali (Indo-European, Bangladesh, India and elsewhere), Chukchee (Chukchi-Kamchatkan, Northeast Siberia), Hungarian (Uralic), Indonesian, Nahuatl (Uto-Aztecan, Mexico),

[53] The form *ye* is unusual in that it appears medially in this otherwise VSO language.

Swahili (Niger-Congo, Zaíre), Vietnamese (Austro-Asiatic), and Yagaria (Trans-New Guinea, Papua New Guinea). In both Bengali and Nahuatl, the same form that is used with locatives and a progressive form shows up in the past tense for NPs and adjectives. For Bengali, McWhorter provides the examples in (21), from Ferguson 1972.

(21) Bengali:
a. Lokti kerani.
man clerk
'The man is a clerk.'
b. Lokti dhoni.
man wealthy
'The man is wealthy.'
c. Boiti tebile ache.
book table-LOC COP
'The book is on the table.'
d. bolchi < bole + achi
say COP.1sg
'I am saying'

Yet another pattern is evident in some of the languages from Hengeveld's (1992) sample. Hengeveld separates "equative" sentences (those with "referring phrases" as predicates) from "ascriptive" sentences (where the predicate denotes a property, not an individual). Gilyak (Gilyak, Eastern Siberia, Northern Japan), Guaraní (Equatorial-Tucanoan, Paraguay), Ngalakan (Gunwinyguan, Northern Australia), and Thai (Daic) allow equative sentences to be copulaless, but no others. Indeed, Hengeveld argues that equative sentences are the most conducive to copulalessness. With the exception of Jamaican Creole and Babungo, any language in his sample that allows copula absence in some environments allows copulaless equative sentences. This would seem to be in stark contrast to the AAVE pattern, as equatives (prototypically, at least) involve NP predicates. However, as the variationist studies of AAVE and related varieties have not typically considered equatives separately from predicative NPs. Therefore, it could be that equatives favor copula absence in AAVE, but are so infrequent as to have little effect on the overall distribution of copula absence before NP.

The remaining languages in the sample (Ewe, Finnish, Gã, Hindi and Yoruba) have categorical copula presence with NP, adjectival and locative predicates. However, the form of the copula can differ across these environments.

As Kiparsky (1988) argues, any universal functional consideration can potentially be overridden or otherwised disguised in particular languages. Nonetheless, this preliminary review of crosslinguistic patterns of copula presence/absence is not promising for a functional explanation of the pattern of non-categorical constraints in AAVE. Above, I suggested two possible functional explanations for the possible similarities and differences across the datasets given in Table 5.18: a functional constraint active in all languages but disguised in some, or a functional constraint that got grammaticized in the development of the varieties studied. Given the cross-linguistic evidence reviewed here, it seems unlikely that a general functional constraint of the type proposed is involved.

Perhaps the root of the similarities is to be found in the similar histories of the varieties studied. Indeed, Rickford suggests that the similarities are due to sharing a history of decreolization:

> Overall, if one simply compares the quantitative patterns of copula absence by following environment in the creole varieties and in AAVE, one is struck by the parallels between them (with one or two exceptions), and it is this parallelism which has provided one of the main planks for the hypothesis that AAVE might have been the diachronic outcome of a decreolization or variation process similar to that synchronically evidenced in the Caribeean, the Sea Islands, and Liberia. (1998:179)

Singler (1991) and Winford (1992) sketch a description of how the process of decreolization might lead to the pattern of copula absence found in modern AAVE. This model is illustrated in Figure 5.9, which is Rickford's (1998:179) adaptation of Winford's Figure 6 (1992:48). The key point is that basilects tend to have different

	Basilect		Lower mesolect		Upper mesolect		Acrolect
NP	*a*	→	Invariant *is*	→	*is*/forms of *be*	→	Inflect. *be*
Adj	ϕ	→	ϕ	→	ϕ/forms of *be*	→	Inflect. *be*
Loc	*de*	→	ϕ	→	ϕ/forms of *be*	→	Inflect. *be*
Prog	*a* V	→	ϕV+*in*	→	(*be*)V+*in*	→	*be*V+*in*
Fut	*a go* V	→	ϕ *goin*+V	→	(*be*) *goin to* V	→	*be goin to* V

Figure 5.9: Model of decreolization in the Caribbean English Creole copula system

copulas for NP and Locative predicates, and no copula with adjectival predicates, as adjectives are just stative verbs in these varieties. Progressive and future forms are again treated separately.[54] Moving up the creole continuum, the basilectal forms are gradually replaced in all environments by forms of *be*. The differences in copula absence across different grammatical environments in the mesolects (AAVE is thought to correspond to a mesolect) correspond to how early forms of *be* come in. This model was constructed on the basis of a study of basilects, mesolects and acrolects (in Winford's case, of Trinidadian Creole). However, as Rickford points out, it is merely a description of what (may have) happened, and provides no explanation for why forms of *be* come in in the different environments at different times.

If AAVE is the result of decreolization, then it shares many things with the creoles Rickford compared it to: the process of creole genesis, the process of decreolization and/or the early establishment of a creole continuum, the fact of being a creole with an English lexifier, and (with the exception of Hawaiian Creole English) the fact of having African substrate languages. The pattern of non-categorical constraints on copula absence might be the result of any one of these factors, or of some combination of them. While functional considerations might clearly be active in the processes of creole formation and decreolization, this is distinct from claiming that functional processes are what determines synchronically the distribution of copula absence across the environments.

[54]Although in Winford's illustration, the future and progressive both involve a form that is homophonous with the copula for NP predicates.

5.4.4 Summary

In this section, I have argued that a functionalist explanation of non-categorical constraints on copula absence is fully compatible with stylistic variation, with copula absence having social value, and with non-categorical constraints affecting that social value. However, the available evidence on copula absence in AAVE, related varieties, creoles, and other languages suggests that any Kiparskian functional constraint in operation is masked in many other languages by interacting factors. This makes establishing such an explanation more difficult, as either the interacting factors would have to be enumerated, or some other means of establishing the grammar-independence of the functional constraint proposed would have to be found. On the other hand, the similarity between AAVE and the other varieties in Table 5.18 could also be explained by similar development. While functional constraints can certainly play a role in the development of creoles and other varieties, this is not the same as using functional constraints in a synchronic explanation: the pattern resulting from the historical development is still a matter of grammar synchronically. At any rate, I will argue in the next chapter that the results of the experiment reported in Chapter 4 entail that the non-categorical constraints are represented in the grammar, even if they are functionally motivated.

Finally, as noted in Section 5.4.1 above, the perception data reported in Chapter 4 suggest that speakers have knowledge of the non-categorical constraint tested, even if the production data are consistent with a functionalist explanation. According to the functionalist account how social value can be associated with individual instances (§5.4.1), in order to rate copula presence as sounding more 'educated' or 'reliable' in the disfavoring environment (__V+ing), the listeners in my experiment had to know that __V+ing is a disfavoring environment for copula presence. This knowledge could either be frequentistic (knowledge of the distribution of copula presence across environments) or qualitative (some direct representation of the functional pressure underlying the distribution), but it must be represented somehow.

If people have and use this knowledge as listeners, they may also use it as speakers, choosing copula absence or presence on the basis of how they expect their listeners to evaluate it in a certain environment. This suggests a way for functionalist constraints on variation to become grammaticized, if the variable in question is socially meaningful. Speakers who know that copula presence before a verb is uncommon, and therefore particularly charged in its social value, might avoid copula presence before verbs except when they do really want to express that social value. Since copula presence before nouns is (known to be) more common, and thus is less charged with social value, it will be used more often. In this way, the system is perpetuated doubly: by whatever functional constraint is operative, and by speakers using the forms differentially because of their social value. Further, this feedback loop mediated by social value could perpetuate a pattern even if its original motivation disappears, as in the decreolization model.

5.5 Conclusion: Social value and grammar

This chapter has considered three previous approaches to modeling non-categorical constraints on variation, with special attention to how they could handle meaningful variation. The three approaches contrast in how much they stipulate with respect to non-categorical constraints in the grammar. They also constrast in how well they can handle the social signficance of variation. One version of the OT approach (Boersma and Hayes 1999) can marginally handle stylistic variation, but no OT approch can handle the other aspects of socially meaningful variation. The variable rule approach can likewise handle stylistic variation, but not the rest. Only the functionalist approach provides a succesful (in principle) model of the social signficance of variation.

The strength of the functionalist account is that it allows for the association of social value with individual forms, because it does not (necessarily) treat variation as inherently probabilistic. On this account, the effect of the grammatical environment on the social value of a variant would follow from Gricean principles: if the speaker went out of his/her way to use a variant disfavored by the environment, then s/he must be particularly interested in conveying the social value associated with that variant. However, this explanation requires that the listener have and exploit either knowledge of the functional constraint in question, or knowledge of distribution of variants across contexts. On this view, stylistic variation or meaning-in-frequency is not the setting of the input probability to "informal" (or by a choice of grammar) but the sum of choices made by speakers over some stretch of time/in some set of interactions.

Chapter 6 discusses the kinds of linguistic knowledge required by this model and whether each should be considered a matter of grammar.

Chapter 6

Conclusion: The boundaries of linguistic competence

The question I began this study with was, How can a competence grammar account for one of the systematic aspects of variation, namely, non-categorical grammatical constraints? The preceding chapters have sought to answer this through a case study of copula absence in AAVE. Chapters 2 and 3 (among other things) elaborate the competence grammar that I am starting with. Chapter 4 presented preliminary evidence that non-categorical constraints are actually a problem for competence grammar: The results of the study suggest that people have knowledge of these constraints, as can be seen by the interaction between non-categorical constraints and social value/social meaning.

Chapter 5 begins with a discussion of the idea that variation doesn't exist solely in a formal plane, but is infused with social meaning, reviewing some of the literature on social meaning and variation. Chapter 5 goes on to ask how existing approaches to non-categorical constraints can handle the social aspects of sociolinguistic variation including the results of the experiment reported in Chapter 4.

In this chapter, I will suggest that the key to a successful account of these facts lies in recasting the notion of linguistic competence. In brief, I will argue that the tighter coupling of competence and performance (or knowledge of language and use of knowledge of language) already adopted in 'performance-plausible' models of competence

grammar (Bresnan 1978, Pollard and Sag 1994, Sag and Wasow 1999) leaves open the possibility of a more inclusive competence grammar. I will propose an account of the effects of non-categorical constraints observed in the experiment that involves adding three types of information to the competence grammar: social value associated with signs, (a certain kind of) probabilistic or at least frequentistic information, and over-specified types. All three extensions have already been proposed in the literature. The resulting model of grammar, not unlike the usage-based model (Langacker 1987, 1990; Kemmer and Israel 1994; Kemmer and Barlow 2000), is a first pass at a model of linguistic knowledge that fits into what Eckert (2000) describes as a social theory of language.

The structure of this chapter is as follows. Section 6.1 considers the question of what grammar is a model of, whether it concerns the language of the community or the language of the individual on the one hand, and whether it should reflect 'externalized' or 'internalized' language on the other. Sections 6.2 through 6.4 discuss the specific types of information that I propose to add to the grammar to account for non-categorical constraints and their interaction with social meaning, reviewing existing proposals along the same lines. Section 6.5 reviews existing criteria for considering phenomena a matter of grammar and proposes some possible alternatives. Finally, Section 6.6 discusses how this fits into a social theory of language (Eckert 2000).

6.1 What grammar is a model of

In order to discuss whether certain phenomena belong 'in the grammar' it is necessary to ground the linguist's grammar by deciding what it is to be a model of. There are two dimensions of possibility: First, a grammar may model the language of a community or the language of an individual. Second, within models of the language of the individual, one may distinguish 'intensional' models that are meant to reflect the actual internalized knowledge of the individual and 'extensional' models that attempt only to generate the set of form/meaning pairs that the individual would accept. The next two subsections treat each of these dimensions in turn.

6.1.1 Community grammar v. individual grammar

Weinreich et al. (1968) argue that the structure in variation is only apparent when the language of the whole community is studied. What looks random and meaningless at the level of the individual speaker, according to them, becomes part of a pattern when that speaker is considered with the other speakers in his/her community. Thus the variable rules that they present are cast as rules of community grammar—partial descriptions of the language of a community.

Chomsky (1986) takes a radically opposed position, arguing that only grammars as theories of the internalized language of individuals (I-language) can have sound scientific basis. A grammar that generates the set of sentences in a language is a theory of that language. But how is a linguist to find out what the relevant set of sentences is? For individual speakers, one could (with infinite time and a speaker of infinite patience) discover the set of sentences accepted by that speaker. Since speakers from the same community don't necessarily accept the same set of sentences, it is not at all clear what set of sentences constitutes the language of a community (E-language). Taking the set of all sentences accepted by at least one speaker in the community means running the risk of including contradictory sentences: pairs of sentences of which any given speaker will accept only one. From a Chomskyan viewpoint, this would potentially obscure parameters of variation. On the other hand, taking the set of all sentences accepted by all of the speakers in the community would surely lead to an impoverished system. Further, as has often been noted, the notion of community is not well-defined; it is not possible to decide exactly which speakers constitute a 'speech community'. Such a listing would be required, in principle, to enumerate sentences in the way imagined here. Chomsky concludes that linguists should be concerned with the language of individuals, since these systems at least are well-defined:

> Theories of E-language, if sensible at all, have some different and more obscure status [than theories of I-language] because there is no corresponding thing in the world. (1986:27)

The thing-in-the-world that corresponds to a theory of I-language is an individual's internalized knowledge of the language. That is, the underlying system behind the patterns of grammaticality is the individual's own cognitive state. There is no cognitive state that contains an underlying system for the language of the whole community, no community mind for the community grammar to be an aspect of.

Part of the issue here is that Chomsky and Weinreich et al. differ in more than where they take the structure of language to reside. Chomsky defines language in terms of grammaticality, and studies the patterns of grammaticality across different forms. Weinreich et al. are studying differences in form that do not affect grammaticality, and are investigating how the patterns of use of the variant forms articulate with the structure of society. Thus what looks ill-defined to Chomsky from the point of view of patterns of grammaticality is the only way to study what Weinreich et al. are interested in.

Further, it's simply not true that successful theories of systematic phenomena must be models of some thing-in-the-world that embodies the underlying system. For example, theories of macro-economics describe the systematicity of the combined commercial acts of whole countries. This does not require that there be a national mind which contains some representation of the underlying system.

Thus is might seem that it is possible to study the language of a community, and to posit a grammar of that language, as long as that grammar is not construed as a model of linguistic competence. However, this does not mean that sociolinguistic patterns are strictly outside the purview of competence grammar. Rather, the thing in the world that corresponds to community grammar is individuals' (imperfect) understanding of how other people speak. In this case, it doesn't really matter that speech communities are vague or ill-defined. Every speaker has some well-defined (although ever changing) set of linguistic and social experiences. From these experiences, speakers build their imperfect models of the speech of other people in their social landscape. Through the process of iconization (Irvine and Gal 2000) forms or patterns of usage become (for the speaker) iconic of types of people, personality characteristics, stances, etc. Note that this process involves analysis of or abstraction over both the linguistic and the social aspects of the speaker's experience. The result of this abstraction is that individuals come to see their communities as well-defined.

There is a danger of missing the crucially social aspect of language if one focuses on grammar as a model of linguistic competence. That is, individuals do not construct grammars any which way, but rather on the basis of their experience of language use in their communities. As Harder (2000) argues, individuals' mental representations of the linguistic system involve or invoke a notion of the collectivity. Focusing on lexical semantics, Harder argues that, from the point of view of any individual speaker, the meaning of most words is a fact about the way the world works.[1] I will return to these points in Section 6.2.2 below. At the same time, it is important not to overemphasize the role of the actual patterns of language in the community. That is, the fact that individual speakers represent their knowledge of language in terms of the expectations of their community doesn't entail that individuals are correct in their understanding of the expectations of the community.

6.1.2 Linguistic competence and psychological reality

In the preceding discussion, I have contrasted Chomsky's and Weinreich et al.'s positions along the community/individual dimension. Chomsky's position also represents a strongly cognitivist view:

> Statements about I-language ... are true or false ... The I-language L may be the one used by a speaker, but not the I-language L′, even if the two generate the same class of expressions ... (Chomsky 1986:23)

This is where the distinction between 'intensional' and 'extensional' models of language comes in. One and the same language (defined as a set of sentences) can be described by multiple different grammars. An 'extensional' model of language makes no claim of psychological reality. An 'intensional' model of language purports to be a model of what speakers actually know.

Wasow (1978) and Soames (1984) argue that the methodology of generative grammar can only provide evidence for extensional models. If linguists draw their evidence solely from the set of sentences accepted (grammaticality judgments), then grammars

[1] The exceptions to this are technical terms that people create and provide definitions for.

written on the basis of that evidence can only be equivalent in extension to actual internalized grammars of real speakers. While the language to be described places some degree of constraint on possible grammars describing that, it is clear that this constraint still far underdetermines the grammar. Any similarity between linguists' grammars built on this kind of evidence and actual internalized grammars is a matter of chance.

There is, of course, nothing wrong with an extensional model of language. It is a legitimate object of study. However, when it comes to providing explanations, and not just accounts, of linguistic patterns, extensional and intensional models have different potential sources of explanatory force. The customary sources in mainstream syntax—language acquisition/UG/the LAD in GB and the interfaces with other cognitive systems at PF and LF in Minimalist theory—belong to the domain of intensional models. Extensional models, on the other hand, draw explanatory force from functional and structural considerations.

However, it is not just the methodology of mainstream syntax that removes it from its claimed sources of explanatory force. Seuren writes:

> In the transformational literature, it is customary to claim psychological reality for competence models on the one hand, but to deny, on the other, that they could be taken as process models. (Seuren 1982:4)

That is, mainstream competence models are supposed to be models of actual linguistic knowledge, but not of the knowledge that gets used in language processing. This separation entails that psycholinguistic results and theoretical considerations about language processing cannot inform the design of grammar, on such a theory.

In contrast, Bresnan (1978), Pollard and Sag (1994) and Sag and Wasow (1999) (among others) argue for what Sag and Wasow call 'performance-plausible competence grammar', that is, models of linguistic competence that are embeddable in models of language processing (performance). On such a model, psycholinguistic results can inform the design of grammar as well as specific linguistic analyses. To the extent that this happens, the model of grammar is indeed a model of the knowledge of actual speakers. Bresnan puts it quite succinctly:

> If a given model of grammar cannot be successfully realized within a model of language use, it may be because it is psychologically unrealistic in significant respects and therefore inadequate in those respects as an empirical theory of the human faculty of language. (1978:2)

Pollard and Sag (1994) motivate their declarative, constraint-based framework by citing psycholinguistic evidence that language processing is highly incremental and highly integrated with information about the world, the context, and shared knowledge between speakers, and that linguistic information functions to like effect in different processing activities (notably, comprehension and production). Similarly, as noted in Section 2.3.4 of Chapter 2, there is psycholinguistic evidence to support a traceless theory of extraction (Pickering and Barry 1991, Sag 1998).

One thing that performance-plausible competence models of HPSG and LFG share with mainstream syntax (GB and Minimalism) is their general conception of what belongs in the grammar. Chomsky (1955) applied the notion, developed by himself and others in Formal Language Theory, of language as a set of sentences and laid out a research program of building grammars that generate the appropriate strings. In this early work, the cognitivist position is not evident. Unconcerned with issues of linguistic knowledge, Chomsky is free to define the object of study in whatever way is convenient. Theories of linguistic competence, including the performance-plausible competence models discussed here, inherited this definition of the object modeled by a grammar. However, theories concerned with actual speakers' knowledge of language do not have the freedom to arbitrarily define the object of study. To my knowledge, there is no evidence that speakers possess a separate module of knowledge limited to that which is usually included in competence grammar.

To put it somewhat differently, I see performance-plausible competence models as modeling the knowledge of language that gets used in language processing. The studies in psycholinguistics and computational linguistics cited below as well as the results reported in Chapter 4 of this dissertation suggest that people have knowledge of language that is of a type not usually included in grammaticality-focused competence grammars. This leaves two possibilities: either people's knowledge of language is partitioned into two parts (grammar and other) or it forms a unified whole. If

knowledge of language is partitioned, there are several lines along which that partition could occur. In the following three sections, I will discuss three different kinds of knowledge of language that I suggest belong on the grammar side of any partition. In Section 6.5, I will discuss a few possible heuristics for determining what is in the grammar and what is not.

6.2 What's in the grammar I: Social meaning

The first additional kind of linguistic knowledge I would like to consider is 'social meaning'.[2] This was discussed to some extent in Section 5.1.2 of Chapter 5, where I argued that social meaning could reside in both individual forms and in patterns of usage. Here, I would like to discuss the content of social meaning—that is, what do socially meaningful linguistic forms *mean?*—and review some previous discussion in the literature about including social meaning in the grammar.

6.2.1 Kinds of social meaning

In order to focus this brief discussion of social meaning, I will organize it around four types of socially meaningful forms:

- Forms associated with some property
- Forms associated with some stance (polite/impolite)
- Forms associated with some situation type
- Forms like *Please*, *Good morning*, *uh huh*

These types are not necessarily clearly distinct from each other, and are not intended as an exhaustive list of socially meaningful forms. Nonetheless, they should serve to organize what follows.

[2]Some would object to this term on the grounds that all meanings are social. Although I can hardly disagree with this point, for lack of a better term, I will continue to use 'social meaning' and contrast it with 'denotational meaning', for ease of exposition.

A potential example of the first type is the fronting of (o) and (oh) in Eckert's (2000) study of the speech of high school students in the suburbs of Detroit. Eckert found that burnouts tend to lead in variables involved in changes that are more advanced the closer one gets to the urban center. Regarding this pattern, she writes:

> This could be taken as an indication of greater access to these changes stemming from the burnouts' urban-oriented life style. It can also be taken as an indication that these urban-led changes have urban-related social meaning that gives them positive symbolic value for burnouts and negative symbolic value for jocks. (2000:136)

Thus these variables have a social value that involves properties associated with 'urbanness'. Using such variables in the construction of one's own style amounts to attempting to sound like one has those properties. Another example of this same type is Ochs's (1992) example of the sentence final particles *ze* and *wa* in Japanese (discussed in Section 5.1.2 of Chapter 5). Both particles express 'intensity' (with regard to the denotational content of the utterance) but they contrast in that *ze* is associated with the property of being coarse, while *wa* is associated with the property of being gentle. Note that in this case, because of the process of iconization, a sincere use of *ze* counts as being coarse, not just as sounding coarse. A similar effect may obtain with the urban variables in Eckert's study.

The association of forms with properties of individuals may be mediated by iconic individuals in the speaker's social landscape. Eckert describes Judy, a "burned-out burnout" as "a social and linguistic icon—a local personage whose extreme embodiment of burnout practice and style serves as a benchmark of social meaning for her cohort" (2000:2).

I suspect that the social value of copula presence/absence is of this first type: it is associated with properties of individuals that speakers may then use in the construction of their own personal styles, etc.

The second type of social meaning involves forms associated with a stance, typically a stance vis à vis the addressee. Thus addressee-oriented honorifics as well as solidarity forms fall into this class. The boundary between this type and the last is not clear cut. For example, the use of Japanese addressee-oriented honorifics not only

expresses a certain relationship between speaker and hearer (i.e., indicates a stance), but may also be associated with certain (non-relational) properties of the speaker. For example, extensive use of all kinds of honorifics is associated with the speech style of middle-aged, upper middle-class women. Further, since addressee honorifics construe the addressee as socially superior, they are also potentially associated with a kind of powerlessness on the part of the speaker. In these latter functions, Japanese honorifics fall into the first type. However, one or more of these 'social meanings' may be indirectly indexed (cf. Ochs 1992). That is, if the meaning of honorifics is a stance of humbleness, their association with powerlessness might be a matter of inference from the repeated expression of a humble stance.

The third type of social meaning involves forms associated with particular situation types. For example, in Japanese, honorifics are used in certain stylized, formal situations, such as the tea ceremony. Their use in this case is independent of the actual social relationships of the participants. Or consider the speech style of livestock auctioneers in the US. This style is associated with the role auctioneer and therefore with the situation in which that role exists (the auction). Similarly, null objects in English are associated with the giving of instructions, as in recipes. I have argued elsewhere (Bender 1999) that stylistic markers like null objects are part of what makes a recipe a *recipe* and not merely a description of the making of something. Thus socially meaningful forms are involved in the social construction of these situation types (tea ceremony, auction, recipe) in at least two ways: On the one hand, they contribute to the texture of the situation. The tea ceremony is an extremely formal event, partially because of the formal language used. On the other hand, they can become a part of the definition of the situation. One hasn't properly performed the tea ceremony if one has used ordinary, non-honorific language during the event. Likewise, a livestock auction certainly isn't prototypical unless the auctioneer spoke in that particular rapid style.

Here again, the boundary between this and the other categories is not sharp. One use of honorifics in Japanese is for 'beautification' (*bikago*). In this use, the honorifics do not serve (primarily) to express politeness or formality, but rather to express a property of the speaker (see Hendry 1992 and the references cited there). Their meaning in this use is, of course, textured and complex, but part of the meaning may be achieved by linking the speaker to the elite, refined situations in which honorifics are used.

I have been considering here individual forms associated with particular situations, but of course more than one form may be associated with the same situation. Indeed, the auctioneer style mentioned above is not a matter of one feature (say, speed of delivery) but rather a whole complex including prosodic and segmental features as well as formulaic phrases (*going once...*). Such complexes constitute registers. In a theory that recognizes socially meaningful forms, one can say that the sum of the socially meaningful linguistic choices in a text (particularly those that relate to situation type) together constitute the register of that text. To the extent that the speaker's linguistic choices are influenced by the situation in the same way, there will be a certain consistency to those choices. As such, in recurrent situation types, similar patterns of linguistic choices may tend to recur, and in time the those patterns may be reified as recognizable, conventional registers.[3] In other cases, the register (in the first sense, as the sum of the linguistic choices) may be more free-form, as, for example, when a politician interrupts a stump speech with a joke, or a side comment.

The fourth type of social meaning is that attached to forms like *please*, *thank you*, *good-bye* and *uh-huh* and similar forms used to provide feedback during another speaker's turn. These forms are also associated with situations, but of a different order.[4] For example, one says *good-bye* upon parting. What is most interesting about these forms is that, while the previous types mostly concern alternations in the form of otherwise semantically contentful signs, this type involves words that have little or no other meaning. The existence of such forms provides the first argument that social meaning is a matter of grammar. As Hudson (1996) argues, to exclude social meaning from the grammar would be to deny these forms the status of signs.

[3]I leave the question of whether knowledge of these patterns should be integrated with other linguistic knowledge, and if so, how, to future work.

[4]The recipe as a situation may fall somewhere between these and the livestock auction/tea ceremony type of situation.

6.2.2 Some complications

The preceding remarks only scratch the surface of the issues around social meaning. A thorough discussion is beyond of the scope of this study. Rather, my purpose here has been to sketch the kinds of phenomena that I believe require some reflection in competence grammar. But what, exactly, should go into the grammar? Clearly, the meaning of Japanese honorifics does not include complete information about the tea ceremony, or about the particulars of the various relationships that the speaker finds him/herself in. Rather, what's in the grammar is probably something more abstract, some collection of relatively underspecified social meanings that can be used to various effects in situated speech.

Determining what those meanings are for AAVE copula absence/presence is beyond the scope of this study. The ethnographic methodology that Okushi (1997) developed and applied to the study of Japanese honorifics would seem to be a good starting point, however. Extrapolating from her study of the use of honorifics, one might posit a set of social meanings for these forms which includes 'formality', 'politeness' and 'distance'.[5] Thus the honorifics are used in the tea ceremony for their value in expressing the formality of the occasion, and the representation of their social meaning in the grammar need not include any reference to the tea ceremony. In Morgan's (1978) terms, the fact that honorifics get used in the tea ceremony is a convention of usage (and an extragrammatical fact about the tea ceremony), while the fact that honorifics can express formality is a convention of language. However, this does not mean that the use of honorifics in the tea ceremony is irrelevant to their social value. Their use in formal situations such as the tea ceremony reinforces (reproduces) their social value as formal. Another way in which the meaning-in-use of Japanese honorifics goes beyond their basic meaning is in what Okushi terms 'playful' or 'sarcastic' uses of honorifics in situations that don't warrant them. That is, speakers can achieve a playful or sarcastic tone by playing off the normal expectations for the use of honorifics.

Note that in this respect social meanings are similar to denotational or referential

[5]These are clearly related stances, and may be reducible to one underlying meaning, a question I leave to future research.

meanings. Whether or not one includes encyclopedic information in semantics (and therefore in the grammar) (cf. Langacker 1987), the meanings of ordinary words get vivified in everyday use (Sag 1992). For example, consider the pair *hear* and *listen*, which mean something like 'perceive aurally' and 'pay attention to what one perceives aurally'. In Mandarin, there is the pair *tīng* and *tīng jiàn* with similar meanings, but in this case the contrast has to do with successful perception. *Tīng* means something like 'listen for' and *tīngjiàn* something like 'succeed in hearing'. Now consider the following short discourses in English:

(22) A. Did you hear what I just said?
B. Sorry, I heard you but I wasn't listening.

(23) A. Are you listening to me?
B. Yes, but I didn't hear you until just now.

In (22), the notion of 'successful perception' appears to be associated with *listening*. In (23), it appears to be associated with *hear*. In other words, 'successful perception' is not really associated with either word in English, and depending on how their meanings are vivified in context, either can include it.

Sag (1992) provides the example in (24).

(24) Craig cut the lawn/hair/cocaine/record/rookie.

Here the meaning of *cut* is quite different depending on the object. In at least the first three cases, the sense of *cut* involves separating a single thing or mass into two parts. However, the actual actions that Craig engages in in each case differ. On hearing that Craig cut the lawn, we don't imagine that he got down on hands and knees with a pair of scissors and trimmed the blades of grass so as to make the lawn more stylish. Thus denotational meanings, like social meanings, are subject to vivification in actual contexts of use. The fact that the latter senses of *cut* (with *record* and *rookie*) may be seen as metaphorical only shows how powerful this process is.

To summarize the remarks on this first complication, social meaning, like denotational meaning, is subject to vivification in its actual contexts of use. This complicates considerably the task of deciding what exactly about the social meaning of an item

belongs in the grammar, but does not detract from the point that some social meaning does belong.

Another complication that arises with social meaning is the possible specificity to particular groups of interlocutors. That is, it is plausible that copula absence/presence has some social meaning for AAVE speakers that it does not have for other groups, and that AAVE speakers are aware of this. Here again, there is a similarity with denotational meaning. Take any ordinary word that also has a use as a technical term, for example, *daughter* (of a node in a tree), *mouse* (attached to a computer), *government* (of a dependent by a head), etc. Speakers who know the technical meanings of these words probably also know which interlocutors are likely to share that meaning.

At first glance, this information would seem to be an implausible candidate for inclusion in the grammar. However, Harder (1996, 2000) presents a view of linguistic meaning as distilled out of communicative practice, where sentences do not encode messages so much as provide hints for the interlocutor to arrive at that message. Harder writes:

> In this picture, the meaning of the word *horse* is distinct from the concept HORSE per se: the word *horse* is a means of interaction, which invokes the concept HORSE for purposes of interaction—and the meaning consists in [the] act of invoking the concept. (2000:9)

If this view is right, and meaning involves a representation of the effect that using a word can have on an addressee, then such a representation could also specify the type of addressee one should expect to be so affected.

One way in which social meanings do differ from denotational meanings is that one would expect the former to be subject to more disagreement between speakers. That is, while most English speakers probably have fairly similar meanings associated with the word *horse* (or in Harder's terms, similar beliefs about what concept its use evokes), certain kinds of social meaning appear to be much more fluid, more easily molded for local stylistic purposes, and more likely to be misunderstood. This is not to say that denotational meanings are not subject to (re)negotiation, for they patently are. One need only look as far as Clinton's famous "that depends on what the meaning of *is* is." Likewise, words that have undergone semantic change provide evidence of past renegotiation: *will* (once meaning 'wish'), *mistress* (once parallel to *master*), *awesome* (once meaning 'awe-inspiring', now similar to *cool*)... Thus what we have here is not a difference in kind, but rather a difference in degree.[6]

One might think of this difference in terms of degree of conventionality. But what is convention? For Lewis (1969), conventions are norms or agreements to which each participant prefers to conform (as long as everyone else does) rather than being the sole dissenter, and which could have been otherwise. Thus, using the word *horse* to invoke the concept HORSE is a convention because any individual English speaker should prefer to maintain that association as long as everyone else does: dissenting alone would lead to being misunderstood and failing to understand other people, but there is no point in sticking to it if no one else does. Further, the convention could just as well have been otherwise, as indeed it is in other languages. Searle (1995) sees facts about language on a par with other 'institutional' facts about the world. A piece of paper with a certain kind of printing on it counts as a $5 bill, because we collectively assign that function to it. A $5 bill counts a money, because we collectively assign that function to it. Likewise the word *horse* can serve to evoke the concept HORSE because we collectively assign that function to it.

Lewis (1969) (see also Clark (1996)) describes the mutual knowledge involved as a recursive sequence of "you know that I know that you know...". Searle, on the other hand, argues that it is a matter of irreducibly collective intentions. This doesn't require any kind of community mind. Rather, individual speakers hold beliefs that start with "We intend that..." This notion of collective intentions is important because, as Harder (2000) points out, this is what accounts for the fact that, to speakers, the meaning of a word is not a matter of their choice, but rather a fact about the world.

Harder (2000) argues that there is still more to it, however. Conventions of language rely not only on collective intentions but also on communicative practice. It is indeed communicative practice that makes facts about language facts about the

[6]Changes in denotational meaning are often highly tied up with aspects of social meaning, and changes in both kinds of meaning only take place in socially situated interaction. This is particularly clear in the case of words like *mistress* (McConnell-Ginet 1984, 1989).

world. Harder offers the analogy of traffic lights:

> [T]he function of traffic lights would not change merely because an observer decides to withdraw his allegiance—his only chance would be to try to disrupt the actual pattern of road interaction in such a way that the causal feedback cycle itself was threatened. (2000:11)

For Harder, the causal feedback cycle is the defining feature of functions. Something has a function if it forms a part of that larger whole and it is causally implicated in the preservation or reproduction of the larger whole. In the case of language, the larger whole is the internalized systems of individual speakers. The usefulness or function of any part of that system is dependent on the speakers' experience with communicative practice up to that point. Likewise, any given use of language provides input to the next stage of the interlocutors systems. That *horse* can be used to evoke the concept HORSE among English speakers is a fact about the world. Humpty Dumpty deciding otherwise won't change that. However, as with traffic lights, Humpty Dumpty could try to disrupt or change the feedback cycle by consistently using *horse* to refer to something else. This doesn't guarantee a change, but could lead to one if Humpty Dumpty has the right social status, etc.[7,8]

So how does this apply to the difference in malleability between social and denotational meaning? One thing to note is that it is possible for speakers from different English-speaking communities to have a conversation in which the denotational semantics functions as intended without understanding the social meaning that may have also been expressed. For example, I could go to Detroit (where the vowel shift has been extensively studied by sociolinguists) and ask for and understand directions without understanding the social meanings associated with my interlocutor's rendition of the vocalic variables. The effect of this interaction on my grammar might

[7]To get *horse* to stop meaning HORSE, Humpty Dumpty would also have to introduce a new form for HORSE and hope that that catches on as well.

[8]This view of convention as rooted in practice is echoed by Eckert:

> Convention is not a thing but a process, and the possibility of convention resides in speakers' ability to hypothesize about others' behavior and to take interpretable action, along with a commitment to doing so within a particular social unit. (2000:45)

include improved recognition for words spoken with a Detroit accent, as well as an association between that accent and whatever salient properties I perceived in my interlocutor (perhaps just the fact of being from Detroit, perhaps something more related to personality characteristics). However, I would have to get to know my interlocutor and her community much better before I would learn anything about the social meaning attached to those variables for her.[9]

This is just another way of saying that we expect the broader English speaking community to share a meaning for e.g., *turn right*, but not the social meaning of variable vowel pronunciations. Where might this expectation come from? I think in this regard one cannot ignore the influence of language standardization—schooling, dictionaries, widely available written sources—in shaping those expectations. But there may well be more to it, in particular, the kind of concepts that the different kinds of meaning refer to. Why should Belten High students expect an outsider to understand the social meaning of (ay) raising, if that social meaning for them is based in a shared social landscape populated by individuals like Judy (the burned-out burnout, the social icon) whom the outsider doesn't know? On the other hand, I can expect any adult to be familiar with the notion of turning right. The in-between cases also fall into line here. Politeness is a kind of social meaning that I expect most adult speakers to know about, and indeed Brown and Levinson (1987) are able to isolate universal aspects of politeness systems. I would suspect, then, that associations between form and politeness-based social meaning would remain constant across larger groups of speakers than other kinds of social meaning. Similarly, I don't expect all adult speakers to be familiar with the concept evoked by *c-command*, and technical terms have currency only in smaller communities of use.[10]

Following this line of thought, fluidity is the flip side of locality. It is easier to change local meanings than more global ones because local meanings have a smaller mass of expectations supporting them. One way to look at it is that, with a local meaning, each individual speaker accounts for more of her interlocutors' experience

[9]Having heard my out-of-town accent, my interlocutor might have surmised all this and it may have had an effect on her performance.

[10]The phenomenon of slang would be particularly interesting to investigate from this perspective, as part of its function is to deliberately limit comprehensibility to members of the in-group.

with the item in question. Because of the kinds of concepts they involve, social meanings tend to be more local and thus more fluid. However, social meanings can be global and denotational meanings can be local, and I would predict that their relative fluidity would depend more on their locality than on the type of meaning involved.

To summarize, this section has made the following suggestions about social meaning: Social meanings boil down to expectations about the reaction of others to certain linguistic forms, and these expectations are grounded in actual practice and individual speakers' experience with that practice. Social meanings in the grammar are somewhat abstract or underspecified and can get vivified differently depending on the situation of use. The social meaning recorded for a given item can be relativized to different types of addressee. The first property is likely to be equally true for denotational meanings. As for the other two, social meanings differ from denotational meanings only in degree: social meanings may be more fluid and more subject to vivification than denotational meanings. However, it might well be that there is no clear distinction between social and denotational meanings along these dimensions but that rather they merge into each other (see again McConnell-Ginet 1984, 1989.).

6.2.3 Arguments for inclusion in the grammar

I am not the first to suggest that social meanings constitute an aspect of signs that should be included in the grammar. In this section, I will briefly review proposals and arguments in the literature.

The first argument, mentioned already above, concerns words that appear to only have social meaning (e.g., *please*, *good-bye*, *uh-huh*) (Hudson 1996:257). If signs are pairings of form and meaning, the only meaning available here is the social meaning. Exclude it, and the status of these words becomes obscure indeed.

Likewise, certain contrasts in form correspond only to contrasts in social meaning. A clear case of this is verb-form honorifics in Korean and Japanese. On Pollard and Sag's (1994) analysis of Korean referent honorifics, the contrast in form corresponds to a contrast in the presuppositions. That is, a verb with a honorific ending introduces an *owe-honor* relation into the BACKGROUND. This last is a feature within CONTEXT that, among other things, encodes the presuppositions associated with using a given sign. The manner in which such presuppositions introduced by lexemes combine to give the presuppositions of the sentence remains to be worked out, but as properties of signs, they are a matter of grammar.

I believe that the results of the experiment reported in Chapter 4 suggest that the social meaning of copula absence/presence in AAVE is an entity similar to the social meaning of honorifics in Korean or Japanese. That is, the contrast in form corresponds to a contrast in social meaning. The obvious difference is that, like much denotational meaning, the proper use of honorifics is a matter of formal study in these cultures. In this respect, the meaning of honorifics can be expected to behave more like a denotational meaning in its currency among large groups of people and in its inertia with respect to local renegotiation. However, I conjecture that a corpus-study of honorifics in either language would turn up internal constraints that, in turn, interact with the social meaning much like what I found in AAVE.

A second argument comes from the idea that the conventions of language are rooted in communicative practice. Given that social meanings are part of our experience of language, why should we expect them to be stripped away? On this point, Langacker writes:

> [U]nits are acquired through a process of **decontextualization** ... If a property (e.g., the relative social status of speaker and hearer) is constant to the context whenever an expression is used, the property may survive the decontextualization process and remain a semantic specification of the resultant unit. (1987:63)

This argument presupposes a view of language acquisition that is at odds with that of Principles and Parameters theory. Rather than seeing language acquisition as a matter of discrete parameter setting in response to triggers in the language learner's experience, the view presupposed here emphasizes gradual accumulation of knowledge of language through experience and generalization over that experience. However, regardless of the theory of acquisition of syntax, acquisition of lexical knowledge—detailed and language-specific as it is—must follow more of a gradual accumulation

model. Thus this argument should go through at least for social meaning attached to individual lexical items, even if syntax is acquired more discretely.

Finally, Hudson (1996) notes that facts about who uses a form refer to the same thing (the same *sign*) as other facts about that form (e.g., its meaning, its part of speech) and wonders why the former type of fact should be treated separately. Of course, it is possible for different kinds of knowledge to refer to the same thing and yet remain distinct types of knowledge (e.g., encyclopedic and semantic knowledge both make reference to some of the same concepts). However, I believe that social meaning is linguistic in the relevant sense, and that heuristics for determining what is and isn't in the grammar will class it with other facts about signs. Some possible heuristics are discussed in Section 6.5 below.

6.3 What's in the grammar II: Minimalist v. maximalist syntax

In the previous section I argued that social meaning can attach to signs and therefore should be considered a matter of grammar. In the case of AAVE copula absence, social meaning appears to be attaching to a lexical sign—(overt) forms of the copula—but with differential intensity depending on certain properties of the complement of the copula. Now, as argued in Section 2.5 of Chapter 2, the distributional evidence supports positing entries for the copula that are underspecified with respect to the part of speech of the complement. This gives us two options: a general entry with an intensity of social value that is a function of the part of speech of the predicate, or more specific entities that specify the part of speech of the predicate and the intensity of the social value. Because of the relationship of the intensity to frequency (discussed in Section 6.4 below) the second choice appears to be more promising. This section briefly reviews evidence in the literature regarding other phenomena that seem to require the addition of extra, overspecified types to the grammar.

6.3.1 Usage-based model

Langacker (1987, 1990, 2000) proposed a model of linguistic knowledge that he calls the usage-based model (see also Kemmer and Barlow 2000). This is a sign-based model in that all linguistic units are pairings of form and meaning. On this model, such units are extracted by generalizating over experience and then recursively generalizing over the units extracted. Langacker writes:

> The grammar lists the full set of particular statements representing a speaker's grasp of linguistic convention, including those subsumed by general statements. ... Speakers do not necessarily forget the forms they already know once the rule is extracted, nor does the rule preclude their learning additional forms as established units. Consequently, particular statements (specific forms) coexist with general statements (rules accounting for those forms) in a speaker's representation of linguistic convention, which incorporates a huge inventory of specific forms learned as units (conventional expressions). Out of this sea of particularity speakers extract whatever generalizations they can. (1987:46)

This is what might be called a maximalist model of grammar. In HPSG terms, a minimalist grammar would include only those types that are necessary to generate all and only the sentences of the language. A maximalist grammar would include further subtypes that are more specialized. Such a conception of grammar goes against the common conception of theoretical parsimony in linguistics, which favors 'rich deductive structure', i.e., maximally simple underlying systems from which all of the data be can produced, although sometimes only with long chains of reasoning. However, as Henderson (1989), Jurafsky (1996) and others have noted, psychological models tend to view storage as relatively cheap and processing as relatively costly.[11] Now, sign-based grammars (including HPSG) already tend towards greater numbers of stored units (compare constructions in sign-based frameworks to the simple, uniform, phrase structure backbone of GB/Minimalism). The usage-based model takes the next step, including (partially) prefabricated pieces as it were: chunks of linguistic structure that could be built up out of smaller pieces already existing in the grammar, but that

[11] See also Bresnan 1978:14.

are frequent enough and useful enough to have around in addition to the minimal types.

Again in HPSG terms, the general picture is as illustrated in Figure 6.1. In this schematic picture, the top part represents the minimal grammar—those types that are absolutely required for descriptive, generative adequacy. However, the more specified types can also fit into the same hierarchy as further specializations of the existing types. Examples are frequently used phrases like *I don't know* (discussed more below), the canonical forms of idioms and collocations (see Riehemann forthcoming), and the complement-specified subtypes of the copula proposed here.

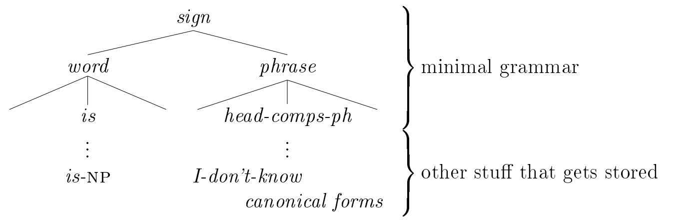


Figure 6.1: Extended grammar

Now, it should be possible in principle to "prune" the hierarchy shown in Figure 6.1 to produce the minimal grammar. However, since all of the information in this hierarchy is represented in a uniform format, it is difficult to see why speakers would partition their knowledge of language along that pruning line. Thus if speakers can be shown to have knowledge of more specific types, as is predicted by the usage-based model, then there is reason to believe that that knowledge is a matter of grammar. The next subsection explores some evidence that speakers have this kind of knowledge.

6.3.2 Linguistic evidence

The evidence for storage of extra types falls into three general classes: the existence of collocations, distributional evidence, and evidence from variation across speakers. Each will be treated briefly in turn here.

To begin with the question of collocations, it would seem that their very existence is evidence of storage of redundant types in the grammar. Examples include *avoid* x *like the plague*, *everything you always wanted to know about* x *but were afraid to ask*, and *get up on the wrong side of the bed*. These phrases all sound familiar, and indeed are listed in collocation dictionaries.[12] Further, phrases that are similar to but not the same as these collocations sound odd, or at least distinctly less typical. Compare, for example, (25) and (26).

(25) get out of bed on the wrong side

(26) avoid x like a tax audit

(25) is actually the form that Kirkpatrick lists for this collocation, but it sounds distinctly wrong to me. Perhaps this collocation is subject to regional variation in form (Kirkpatrick lives in Scotland), but the important point is that while (25) certainly means the same thing as *get up on the wrong side of the bed*, it is possible to recognize one and not the other as the usual way of saying it. Similarly, (26)[13] contrasts with *avoid* x *like the plague* in that the former sounds fresh while the latter sounds clichéd. Anyone making such a judgment must have a representation of these phrasal chunks. But such a representation would be redundant in a minimal grammar because the phrases can certainly be built up out of smaller pieces.

An example of distributional evidence for extra types comes from Bybee and Scheibman's (1999) production study of the reduction of *don't* to /ɾõ/ or /ə̃/. They found that this reduction occurred by far the most frequently in common, fixed phrases containing *don't*, such as *I don't know* or *why don't you*. From a strictly distributional point of view, this indicates that these phrases must be stored as linguistic units in order to be the domain of application of the *don't* reduction rule. Further, if the common phrases are stored as units, the fact that *don't* reduction occurs here and not elsewhere (or, eventually, here more than elsewhere) acquires a natural explanation: If these units constitute single production routines that are recognizable as single chunks by addressees, there is less pressure to articulate each

[12] In particular, Betty Kirkpatrick's *Clichés*, New York: St. Martin's Griffin, 1996.
[13] An authentic example collected by Susanne Riehemann (p.c.).

piece clearly. Further, frequent use of these units on the part of speakers could result in phonetic change that doesn't spread to other instances of the word *don't*.[14]

The evidence from variation across speakers comes from Zwicky's (1982) study of stranded *to*—instances of complementizer *to* separated from its VP complement by a parenthetical or by ellipsis of that complement:[15]

(27) a. I made a decision whether to — and he made a decision, too — have him in the group.
(John Lennon, *Rolling Stone Interviews*, p.150)

b. . . . I can put my dukes up now if I have to in life.
(Joni Mitchell, *Rolling Stone Interviews*, p.389)

Zwicky presents the results of a grammaticality study involving 74 native speakers in which he finds that there is considerable variation in the constraints on stranding *to* in this manner. While all speakers tend to agree on the core examples (i.e., examples of patterns with high frequency in text data), they disagree on the less frequent cases. For example, some, but not all speakers, require the constituent before stranded *to* (the host it leans on phonologically to when its VP is absent), to be either a VP or a predicator within a VP. The examples in (28)[16] don't meet this constraint, and are accepted by only some of the participants in Zwicky's study.

(28) a. %I really need someone to fix the radiator, so it was awfully nice of Helen to.

b. %I realized that you sometimes stay out past midnight, but it seems excessive *always* to.

c. %When they wanted someone to make dinner for 30, they persuaded Quentin to.

Zwicky observes that this 'dialectal' variation does not correlate with social or geographical factors. Instead, he suggests that the analysis of this phenomenon is "underdetermined by the data available to the child acquiring English" (p. 51), and

[14] Further evidence for the existence of such over-specified types comes from the fact that social meaning can attach to them. For example, Penelope Eckert (p.c.) suggests that /aõno/ is a socially meaningful way of pronouncing *I don't know*.
[15] The examples in (27) were collected by Zwicky 1982:9.
[16] From Zwicky 1982:39.

that different speakers generalize from the available data differently. Some end up with very conservative generalizations, in effect requiring all of the properties shared by the frequent, core examples in an example they will judge acceptable. Some generalize further, and in divergent ways.[17] This state of affairs is exactly what one would expect given a model in which speakers slowly abstract generalizations from linguistic experience.

Further, Zwicky notes even speakers who reject the sentences in (28) above accept one class of exceptions to the constraint that rules out (28a–c), namely, sentences in which *to* attaches to monosyllabic function words (29).[18]

(29) a. I want to calculate the bill, but I don't know HOW to.
b. You must write a thank-you note, because not to would be impolite.
c. George didn't do it, until Kim persuaded him to.

By way of explanation for how this state of affairs could have come about, Zwicky writes:

> If a nonlexical item [= function word–EB] frequently occurs in combination with some word, especially a word like *to* with very idiosyncratic distribution, then the combination can easily be interpreted as a unit on its own. Then, when unquestionable combinations of word-plus-*to* become available via the application of [VP ellipsis] after verb-*to* sequences, other closely knit word-*to* sequences are natural targets in an extension of the domain of [*to* stranding]. (1982:45)

These 'word-*to*' sequences are just the kind of prefabricated chunks that are included in a maximalist view of grammar.

6.3.3 Summary

This section has briefly reviewed some theoretical and empirical arguments for including extra, overspecified types in the grammar. One consequence of this is that

[17] There are, however, boundaries to how permissive a grammar speakers will construct. Zwicky argues that these are provided by universal grammar.
[18] (29a–b) are from Zwicky 1982:14,26.

non-local trees might get stored, as in the case of *I don't know* and *why don't you.*[19] This raises certain issues that I will return to in Section 6.5.5 below.

6.4 What's in the grammar III: Probabilities

In this section I will discuss two different kinds of probabilistic or frequentistic information that speakers appear to have knowledge of, postponing arguments about whether one or both belong in the grammar until Section 6.5 below. The first kind corresponds to what might be seen as the resting activation of various linguistic elements (such as alternative word senses, valence patterns, or constructions). This kind of information plays a role in processing, but can be overridden by contextual constraints. The second kind of probabilistic information corresponds to the non-categorical constraints on sociolinguistic variation studied here. It is not immediately clear whether these should be treated in the same way, or if they have differing status.

6.4.1 Resting activation

MacDonald et al. (1994) and Jurafsky (1996) argue that a number of processing phenomena can be best captured by models in which the processor makes use of frequentistic information associated with signs. For example, MacDonald et al. (1994) argue that the relative frequency of different argument structures of verbs can influence the processing of main verb/reduced relative ambiguities. They report the results of an experiment by MacDonald (1994) in which reading times of sentences like (30a–c) were compared.

(30) a. The rancher knew that the nervous cattle *moved* into the crowded pen were afraid of the cowboys.
b. The rancher knew that the nervous cattle *pushed* into the crowded pen were afraid of the cowboys.
c. The rancher knew that the nervous cattle *driven* into the crowded pen were afraid of the cowboys.

[19] Such non-local trees are not, however, necessary for the treatment of copula absence/presence proposed here.

In (30a) and (30b) there is an ambiguity at the italicized word. That word could either be a main verb, taking the NP *nervous cattle* as its subject, or the head of a reduced relative modifying *nervous cattle*. The rest of the sentence provides disambiguating information in favor of the reduced relative reading. Note that in the reduced relative, these forms are passive participles. (30c) is unambiguous because *driven* can only serve as a participle and not as a past tense main verb. (30a) and (30b) differ in that *move* occurs more frequently as an intransitive verb, while *push* occurs more frequently as a transitive. This difference is relevant here because the passive, reduced relative reading requires a transitive argument structure. Thus MacDonald (1994) predicts that (30a) should have more of a garden-path effect than (30b). This prediction is borne out: reading times at the point of disambiguation (*were afraid*) were reliably longer with the *move* type verbs than the unambiguous control sentences (like (30c)), while reading times at that same point with the *push* type verbs do not differ from the control sentences. This result shows an effect of relative frequency of argument structures on processing; that is, it shows that speakers have knowledge of this frequentistic effect and bring it to bear in comprehending sentences. MacDonald et al. (1994) make similar points regarding PP attachment ambiguities and NP/sentential complement ambiguities.

Jurafsky (1996) reports on a computational model of human natural language processing. This model incorporates probabilities attached to constructions and valence patterns. The probabilities associated with each construction and valence pattern used in a parse can be combined to calculate the probability of the entire parse. The parser entertains several possible parses in parallel, but prunes any that fall below some threshold of probability. This pruning is motivated as a means of keeping the computation tractable in the face of the massive ambiguity of human language. With this model, Jurafsky provides a uniform account of frequency effects in the access of lexical items, idioms, and constructions, differing access points in different idioms, attachment preferences (e.g., PP attachment), and garden path effects related to frequency of constructions, valence patterns, and both together.

Now, these frequency effects are such that they can be overridden by context. For example, the strong garden path effect in (31) is partially due to a preference for

raced to appear as in intransitive.[20]

(31) The horse raced past the barn fell.

However, this sentence can be rendered interpretable, and thus the preference can be overridden by an appropriate context:[21]

(32) The horse that was raced down the hill got there just fine. The horse that was raced along the river had some trouble. And the horse raced past the barn fell.

This property of defeasibility or interaction with context suggests a possible distinction between grammatical and other knowledge of language. According to this distinction, grammatical knowledge is that which remains relatively constant, while frequentistic information that can be so radically modified by context belongs to a separate ('performance') component of knowledge. However, here again the distinction is not so clear cut. For example, in the discussion of vivification above, it was noted that the meaning of words also interacts rather dramatically with context. In that case, it seems fairly clear that the underlying, more abstract meaning is a matter of grammar, and the more specific meanings might not even be a matter of knowledge at all; rather they are context-specific effects computed on the fly. If this frequentistic knowledge is analogous, then the resting probability or activation would be a matter of grammar and the actual activation in some context of use computed on the fly. My main purpose in this section, however, has been to introduce this kind of knowledge so that I may ask whether non-categorical constraints belong to the same type.

6.4.2 A tentative formal proposal

In order to consider the status of non-categorical constraints, it will be helpful to spell out a tentative model. This section will do so, building on the work of Jurafsky (1996) and Kemmer and Israel (1994).

Kemmer and Israel (1994) present an analysis of grammatical and social constraints on t/d deletion within the usage-based model. The essential properties of

[20] Other factors include the plausibility of *horse* as the subject of intransitive *race*.

[21] This type of example is due to Crain and Steedman (1985), who are in fact arguing against a probabilistic model.

their analysis (for present purposes) are as follows: All common words subject to variable t/d deletion have two representations in speakers' grammars. The existence of many such pairs leads to the abstraction of a general schema of t/d deletion. However, here we will concentrate on the individual words. In any given pair, say /læs/ and /læst/ for *last* the more frequent alternant will be more entrenched, that is, have higher resting activation. Both forms compete for selection when the speaker goes to pronounce the word, but the more entrenched form tends to win out, in proportion to how entrenched it is. However, both forms may be associated with information about the social situation, etc. If the characteristics of the social situation 'match' the specification on the less frequent item,[22] then the less frequent item will be promoted with respect to the more frequent item, and vice versa.

Note that this approach involves the same kind of random application of probabilities as the OT and Variable Rules models considered in Chapter 5. As such, it is not possible for it to model social value associated with individual instances of linguistic variables. Any given instance of the less likely alternative could either be the result of the random application coming up with the less likely alternative, or of the speaker tilting the scales in favor of that alternative by a desire to express the associated social value. The listener would have no way to tell which it was. Nonetheless, the idea of associating social value with signs, and letting that social value play a role in the selection is promising.

An interesting feature of Jurafsky's (1996) model of parsing is that it uses probabilities in a more deterministic way. Essentially, the probabilities of all of the elements of a parse, resting and context-dependent, are combined to give the overall probability for that parse. The same is done for all of the active parses under consideration, and the most probable parse is chosen. Although the system described in Jurafsky 1996 is a parser, Jurafsky has suggested (p.c.) that the same ideas could apply in production. In this case, given a communicative intent, the speaker selects the means of conveying that intent with the highest probability. On the face of it, this model

[22] Or, more accurately in my opinion, if the speaker wishes to do his/her part in constructing the social situation to match that specification.

would countenance no variation; only the most frequent variant would ever get produced. However if we let the variants be associated with social value, as in Kemmer and Israel's (1994) model and as suggested in Section 6.2 above, then that social value could be part of the communicative intent.

What I have in mind is a system in which the desire to express the social value associated with the disfavored variant can override the bias against that variant. However, the stronger the bias against the element, the more the speaker has to want to express the social value in order to overcome it. Since copula presence is more disfavored with V+*ing* complements than with NP complements, it takes more desire to express the social value of copula presence to use it in the V+*ing* environment. Hearers, knowing this, would react to copula presence accordingly, as they did in the experiment.

There are several complications here. First, it is clear that the probabilities involved are not a once-and-for-all matter, but can be relativized to the interlocutor. The probabilities a speaker uses in production are not necessarily the ones s/he uses in comprehension. A case in point are the non-AAVE speaking African Americans in my study. These speakers showed knowledge of the probabilities of AAVE speakers in their evaluations of the stimuli, but presumably do not apply those frequencies in their own speech. A second complication has to do with the possibility that the same speaker could have different resting probabilities based on who they're talking to, i.e., based on their expectations about how their speech will be heard by that listener.

One prediction of this model is that only the disfavored variant can be used to express social value, a following *ing*-form verb favors copula presence. Interestingly, the results of the experiment seem to indicate that only one variant was meaningful for each group showing an effect (i.e., AAVE speakers and other African Americans). At least, the effect of the following grammatical environment was only apparent for copula presence for the AAVE speakers and copula absence for the other African Americans. However, I don't expect it to be the case that only one variant of any given variable is ever meaningful. I see two other possibilities within this model. The first is that the favored variant is indirectly socially meaningful in that its use indicates a lack of sufficient desire to express the social meaning associated with the disfavored variant. As such, use of the favored variant expresses a weak negation of that social meaning. However, the grammatical environment should have no effect, as it would not be possible to tell how far the speaker's attachment to the social meaning fell short. The second possibility is that other constraints (such as information about the context and the interlocutor) could interact to change which variant is favored, for example, copula presence might be the favored variant when talking to a priest, teacher or other community leader. The grammatical constraints could remain constant in this situation. That is, a following NP favors copula presence more strongly than a following *ing*-form verb. Conversely, a following *ing*-form verb favors copula absence more strongly than a following NP. However, the social value would switch to the newly disfavored variant (copula absence, in this example). Note that these two possibilities (social meaning through negation of the meaning of the other variant, and context-based changes in which is the favored variant) are not incompatible with each other.

The preceding paragraph presumed that there is one, fixed social meaning associated with any variable that speakers can either express with one variant or refrain from expressing by using the other. However, this assumption is not a necessary one for this model of non-categorical constraints. All that is required is that sociolinguistic variables have some social meaning, which may be continually renegotiated. What remains (relatively) constant is the effect of the grammatical environment on the social meaning.

There is a further complication in that non-categorical constraints can in principle be so different or so strong in their effect, that one variant is preferred in one grammatical environment while the other is preferred in another. Indeed, in most of the data sets represented in Table 4.1 of Chapter 4 (see page 144), the V+*ing* environment has a Varbrul weight of over .5, favoring copula absence, while the NP environment has a Varbrul weight of under .5, disfavoring copula absence. However, these weights interact with other factors (e.g., whether or not the subject is a pronoun) and with the input probability, so that it is not possible to conclude from these numbers alone which variant is preferred, and whether that differs across grammatical environments.

For concreteness, I will present a tentative formal representation of my proposal.

However, I regard the preceding discussion as speculative, and offer this formalization only for purposes of illustration. Throughout the discussion, I have been considering four sentence types: copula presence with nominal and verbal predicates, and copula absence with nominal and verbal predicates. However, the grammar doesn't usually provide the choice of nominal or verbal predicates with (roughly) the same meaning,[23] so it is actually forms that differ only in presence/absence of the copula that are competing. This means that the two entries in Figure 6.2 are in competition with each other, as are the two entries in Figure 6.3. In these entries, social value is associated only with copula presence, and indicated by the word 'educated' as a short hand for something that is undoubtedly more complicated. The numbers to the left of the avms are meant to represent differing degrees of resting activation (varying between 0 and 1). It is these values that are used in computing the probability of a parse. The actual values shown in the figure are there only for purposes of illustration, and not based on any particular data set.

$$\textbf{.4}\begin{bmatrix} \textit{copula-be} & \\ \textsc{comps} & \langle \text{ NP } \rangle \\ \textsc{ctxt} \mid \textsc{social} & \textit{`educated'} \end{bmatrix} \quad \textbf{.6}\begin{bmatrix} \textit{silent-copula-ph} & \\ \textsc{args} & \langle \text{ NP } \rangle \end{bmatrix}$$

Figure 6.2: Competing entries for copular sentences with NP predicates

$$\textbf{.2}\begin{bmatrix} \textit{copula-be} & \\ \textsc{comps} & \langle \text{ VP}[\textit{prp}] \rangle \\ \textsc{ctxt} \mid \textsc{social} & \textit{`educated'} \end{bmatrix} \quad \textbf{.8}\begin{bmatrix} \textit{silent-copula-ph} & \\ \textsc{args} & \langle \text{ VP}[\textit{prp}] \rangle \end{bmatrix}$$

Figure 6.3: Competing entries for copular sentences with V+*ing* predicates

According to the model sketched here, all else being equal, copula absence would prevail in both cases. However, should the speaker desire to express the social value of copula presence (e.g., should the speaker want to sound educated), there is the option of using an overt form of the copula. In the case of an NP predicate, the two

[23]Indeed, the test stimuli used in the experiment had to be carefully constructed.

options are relatively close in their activation, and so it doesn't take much to override the resting probabilities. In the case of the V+*ing* predicate, however, there is a much stronger bias towards copula absence and therefore the speaker must be much more emphatic about the social value of copula presence to override this bias. Listeners who have experience with this pattern (and have reason to believe that the speaker is working with the relevant set of probabilities) can interpret this as copula presence with V+*ing* sounding more educated than copula presence with an NP complement.

6.4.3 Interacting factors and context effects

Ivan Sag (p.c.) suggests that one way of partitioning knowledge of language would be into persistent, relatively context-independent knowledge (i.e., grammar) and a component of more context-sensitive, volatile information including the resting activation of various signs. The purpose of this subsection is to consider whether non-categorical constraints are on a par with the kind of resting activation behind garden path sentences. Now, I've just sketched a model of non-categorical constraints in terms of resting activation. However, it is still possible that these constraints interact differently with the context than do other types of bias.[24]

It is first important to note that the experiment reported in Chapter 4 and experiments such as that reported in MacDonald 1994 are not parallel in at least one important respect. MacDonald's (1994) experiment is aimed at investigating how knowledge of probabilities is brought to bear in ambiguity resolution. The experiment reported here, on the other hand, is aimed at investigating how probabilities (or frequency or markedness) affect the social meaning of utterances. That is, the former is a matter of choosing between two alternative structures with their attendant meanings while the latter concerns the meaning of a single structure. When I talk about the speaker's desire to express the social value overriding the grammatical constraint, the latter does not simply disappear but instead has an enhancing effect on the social value.

What about other contextual factors that could interact with the non-categorical

[24] If this is so, a better model of non-categorical constraints might distinguish their implementation from that of other statistical knowledge.

constraints studied here? There are probably numerous factors that enter into the choice between variants in actual production. So far, I have been concerned only with social value and the constraining effect of the grammatical environment (non-categorical constraints). But surely there are also factors, such as speech planning issues, where one might expect to find overt forms of the copula potentially used as fillers in case of disfluency, and recency effects, where a previous use of one form or the other could prime that form for use again. The question is, then, if the non-categorical constraints retain their effect on the social value regardless of these other contextual factors. If they do, then they appear to be distinct from other kinds of statistical knowledge—that is, more persistent, more like other familiar (categorical) grammatical constraints. The resolution of this question will require a more sophisticated model of production and more sensitive experimental techniques.

6.5 The boundaries of grammar: Some heuristics and examples

In the preceding three sections, I have suggested that speakers have knowledge of three kinds of linguistic information not usually included in competence grammars—social meaning, overspecified types, and probabilities of a certain kind—and I explored possible evidence for including this knowledge in grammar. However, by opening up the boundaries of competence grammar to include these kinds of information, I do not mean to imply that grammar is essentially boundless. Clearly, there are kinds of knowledge that are not grammatical. I would like to propose that grammar includes just that knowledge of language that is used in linguistic processing.[25] Presumably, one could devise psycholinguistic tests to determine in any given case whether or not people have knowledge of the information in question and whether they use it in language processing. Here, however, I would like to discuss some general heuristics and in the process sketch a general picture of what I take to be the domain of grammar.

[25] The possibility of partitioning out context-sensitive information will be returned to below.

6.5.1 Arbitrariness/language specificity

One fairly standard criterion for including something in a competence grammar is that of arbitrariness or language specificity (in the sense of specific to a given language). That is, linguistic facts that are conventional in Lewis's sense—they could have been otherwise—must therefore be learned and in the grammar. This criterion is meant to distinguish grammatical facts from natural consequences of functional pressures. It leaks, however, in several respects.

The first problem is it really only works as a sufficient (not necessary) condition for inclusion in the grammar. That is, linguistic facts such as that the word for HORSE in French is *cheval* can be shown by this criterion to be in the grammar, but the non-arbitrariness of other facts does not preclude their grammaticization. That is, just because a pattern could follow from independent principles doesn't mean that speakers don't learn it as a pattern. Indeed, this process is probably what is behind many cases of grammaticization, for example, the development of tones from the different effects of voiced versus voiceless consonants on the formants of vowels (Maddieson 1984). Another example is the tendency to name men before women in when naming couples (Wright and Hay 2000). While this tendency may come from a cultural fact about men being more salient, and thus named first, it is grammaticized in *Mr. and Mrs.* (cf. #*Mrs. and Mr.*) and partially grammaticized in *he or she* (compared to *she or he* which is used with intent to make a point about sexism in language).

The second problem, related to the first, is that language is not as arbitrary as linguists often make it out to be (Langacker 1987). Examples of non-arbitrariness include sound symbolism of the type evidenced by word-sets like *slime, slippery, slink, ...* and morphologically complex words that have drifted just a little from their original semantically compositional meanings. Thus although the full meaning of a word like *fire engine* can not be predicted from its parts, given the meanings of *fire* and *engine*, it is not entirely arbitrary that they should be used together with the meaning of *fire engine*. The same applies for fully compositional words. Langacker writes:

> An obvious but seldom-made observation is that any polymorphemic linguistic sign (this includes the vast majority of expressions) is nonarbitrary to the extent that it is analyzable. For example, given that *staple* means what it does, and that *-er* means what it does, it is anything but arbitrary that *stapler* is the form used in English for a stapling device. (1987:12)

A third problem arises from the process of iconization. If Irvine and Gal (2000) are right, then the connection between form and social meaning is non-arbitrary FOR SPEAKERS CONCERNED, even though it could well have been otherwise. They write:

> *Iconization* involves a transformation of the sign relationship between linguistic features (or varieties) and the social images with which they are linked. Linguistic features that index social groups or activities appear to be iconic representations of them, as if a linguistic feature somehow depicted or displayed a social group's inherent nature or essence. This process entails the attribution of cause and immediate necessity to a connection (between linguistic features and social groups) that may be only historical, contingent, or conventional. (2000:37)

For example, for many English speakers, copula absence and other features of AAVE are iconic of lack of education (or worse, low intelligence) in that they are believed to show or symbolize it directly. Indeed, this kind of social value probably arises as an extension of the stigmatization (by outsiders) of AAVE speakers to the stigmatization AAVE features (again, in the minds of those outsiders). Thus, this kind of social meaning is non-arbitrary in its inception, although it is still arbitrary in the sense that it could have been otherwise. (For example, if AAVE had a different collection of constructions, or if AAVE was not stigmatized but rather exalted.)

To summarize, arbitrariness is at best a one-way criterion for inclusion in the grammar. Many grammatical facts are either not completely arbitrary or not experienced as arbitrary by speakers.

6.5.2 Structure: A system of contrasts

Some authors (e.g., Langacker 1987, Langacker 1990, Hudson 1996) ask, Why should grammar be a separate module of knowledge? They suggest instead a view in which linguistic knowledge is embedded in other knowledge. I think that the motivation for distinguishing (some part of) knowledge of language (the grammar) comes from the fact that the system of contrasts gives internal coherence to grammar, even if it is embedded in a larger knowledge system. One might think of a mosaic, with a white background and the picture picked out in black as an analogy. Although both background and picture are made of tiles, the tiles in the picture together form one coherent system.

This criterion can certainly be interpreted as placing social meaning in the grammar. Social meaning, like denotational meaning, exploits contrasts in form. Further, Irvine (in press) argues that style is a matter of distinctiveness, i.e., contrasts between different groups of people in the social landscape. As building blocks of style, socially meaningful linguistic elements thus pair contrasts in form with contrasts in social space.

One might also be tempted to talk about certain probabilities as a matter of contrasts. That is, copula presence and copula absence are necessarily complementary in their frequency of occurrence because the variable is defined in such a way that choosing one precludes the other. A similar thing could be said for frequencies of valence patterns and the like. However, in this case, it seems that the contrast in frequency merely follows from the contrast in form or in valences. As far as I know, nothing else builds on the contrast in the way that contrasts in social space build on or attach to contrasts in form.

6.5.3 Creativity of use

Another possible criterion is generativity in the sense of creativity of use. Knowledge of syntax and semantics allows speakers to create and understand novel utterances. If other knowledge about language functions in a similar way, by this criterion it should be in the grammar. The California Style Collective (1993) refer to individuals creating their styles as *bricoleurs*, a French word which evokes a picture of building something out of whatever resources and raw materials that one can find. On this view, individuals create their (linguistic) styles by borrowing and combining symbolic

resources made meaningful by their use by and association with others. Socially meaningful forms can thus combine to create novel styles in a manner that is analogous to the way that lexemes and constructions combine to create novel utterances.

It is somewhat less clear how to apply this criterion to non-categorical constraints and overspecified types, but it might be seen to suggest that both belong in the grammar. As for non-categorical constraints, inasmuch as their effect on social meaning allows them to be a tool in the construction of style, they should be considered grammatical for the same reason that social meaning is. With over-specified types, possible evidence for including them by this criterion could come from evidence of people making reference to collocations in the creation of new utterances. Merely using collocations doesn't provide any evidence with respect to this criterion as, by definition, collocations are semantically compositional and can be built up with smaller pieces. However, making reference to collocations means involving them in some sense in the production of novel utterances. An example would be Kraft's marketing of a macaroni and cheese product as *cheese and macaroni*. The unusual placement of *cheese* as the first conjunct is meant to suggest especially cheesy macaroni and cheese. This effect of this suggestion comes partially from the contrast with the usual way of saying it, that is it relies on the knowledge that one usually says *macaroni and cheese*.

6.5.4 Heideggerian *thrownness* of use/acquisition.

Winograd and Flores provide another possible heuristic:

> [T]he essential feature of language activity (the processes of saying and listening) is the thrownness of a person within language. When we are engaged in successful language activity, the conversation is not present-at-hand, as something observed. We are immersed in its unfolding. Its structure becomes visible only when there is some kind of breakdown. (1986:68)

Winograd and Flores define Heideggerian *thrownness* as "the condition of understanding in which our actions find some resonance or effectiveness in the world"

(1986:32). Thrownness contrasts with detached contemplation. Isolated and categorized representations of objects are creatures of detached contemplation. Most language use, however, is entirely transparent. In conversation people pay attention to the messages they are exchanging and not the details of how they are said. Indeed, it takes some training for linguists to learn to pick out examples of phenomena of interest to us from the flow of language use that we encounter in our everyday lives. A "breakdown", in this context, is some kind of problem in the functioning of a tool, such as language. Experiencing a misunderstanding can cause even non-linguists to focus on the words and structures used, to produce and categorize (conscious) representations of the problematic linguistic elements.

By this criterion, the three types of knowledge described in Sections 6.2 to 6.4 above are similar to (other) grammatical knowledge. As Ochs (1992) notes, we learn, comprehend and deploy socially meaningful contrasts in language without paying much attention. That is, it is enough to decide to be informal, or to be polite, or be distant. As competent speakers, we need not pay attention to the individual linguistic choices that implement such decisions. Likewise, we acquired the relevant knowledge with the same seeming effortlessness as the rest of our grammars. Similarly, whatever frequentistic information we've stored is clearly not the result of formal study or conscious tabulation (cf. Haser and Chromiak 1977, discussed in §5.2.1 of Chapter 5). Indeed, part of the reason some authors have doubted that people store frequentistic information is the sheer difficulty of learning it by these other means. Finally, although sayings and proverbs are studied in school in some cultures (notably in Chinese-speaking cultures), this is not the case for the extra types representing such collocations as *avoid like the plague*.

What this criterion does rule out is knowledge of language that is not used in language processing. For example, take the fact that *honcho* in *head honcho* is a borrowing from Japanese. Speakers of English who know this fact about English know it because they were told so explicitly, or because they also speak Japanese and noticed the connection. Both cases involve focusing attention on the word in question, and thus do not constitute *thrown* acquisition. Likewise, it does not seem likely that this knowledge could naturally figure in language processing, and certainly

not without treating it with the kind of detached contemplation that is inconsistent with thrownness. I picked an example from Japanese and not, say, French, because something like the knowledge that *je ne sais quoi* or *raison d'être* are borrowings from French does seem to figure more subtly into their use. French has a certain *cachet*, i.e., air of sophistication, among many English speakers, and a speaker going for this sociolinguistic effect might quite naturally reach for a French phrase. Whether this process is as *thrown* as the rest of language use is open to investigation.

6.5.5 Context sensitivity

The heuristics considered so far seem to converge on the conclusion that social meanings are a matter of grammar. To the extent that they were applicable to the question of extra types, the heuristics also support including the extra types in the grammar. The status of probabilistic information is less clear. In this subsection, I would like to return to the distinction considered in Section 6.4.3, between persistent knowledge that functions to similar effect across contexts and more 'volatile' knowledge that is more subject to modification in context. Modulo the reservations given above about semantics being more sensitive to context than one might suppose, this criterion serves to distinguish knowledge of signs (units of grammar) from probabilistic information about those signs. This subsection briefly describes a potential empirical test that could substantiate this distinction and show that in the minds of actual speakers, these kinds of knowledge constitute separate systems or modules. The test involves finding a situation in which grammar appears to be restricted and then looking for evidence that people have knowledge of language which they use in processing which is not constrained in the same way.

One potential restriction to investigate is that embodied in the HPSG theory of locality of selection (e.g., Pollard and Sag 1994). This theory holds that heads should not be able to select for arbitrary properties of words or phrases contained within their dependents. For example, it predicts that the verb labeled V1 in Figure 6.4 should not be able to constrain the case of the noun phrase labeled NP1. This constraint is enforced by two aspects of the theory: the cancellation of elements of valence lists and the selection of *synsems*. That is, since the complement requirement of the head verb of the embedded sentence in Figure 6.4 is satisfied within the embedded sentence, the embedded S bears the feature specification [COMPS ⟨⟩]. As the embedding verb (V1) selects only the *synsem* of its complement S, it has no access to the DAUGHTERS feature(s) of its complement, and thus no access to information about the object of the embedded sentence. Without this information, it can enforce no constraint on the case or other properties of that NP.

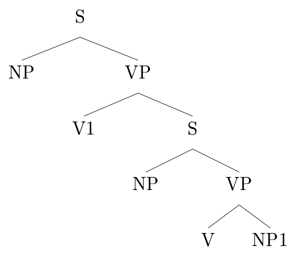


Figure 6.4: Tree illustrating locality of selection

Now, Bod (1998) and Neumann and Flickinger (1999) describe computational systems that make use of statistical information about previously encountered parses to improve disambiguation and parsing performance, respectively. The information recorded by these systems crucially involves non-local trees. By recording such trees, these systems are able to record dependencies that do not obey the hypothesized locality constraint. If people could be shown to record and use the same information, and if the locality hypothesis could be confirmed, then one would have to conclude that grammar forms a separate subsystem of linguistic knowledge.

However, there are several reasons to believe that it might not turn out that way. On the one hand, the hypothesis of locality is maintained in HPSG by using other mechanisms to make available information that is blocked by this constraint. For example, English verbs select for PPs headed by specific prepositions. Given the

locality hypothesis, the identity of the preposition within the PP should be unavailable. To circumvent this, information about the head of the PP can be passed up through semantic features (identifying the preposition by the relation it introduces). However, such prepositions are often semantically vacuous, and so purely formal features are introduced (Pollard and Sag 1987). Similarly, Bender and Flickinger (1999) argue that tag questions attach to sentences (and not VPs). It follows from locality that the information about the subject of the sentence should not be available to the tag modifier. But tags patently agree with the subject, and therefore Bender and Flickinger (1999) take this as evidence that verbs should bear an agreement feature. This feature is within the SYNSEM, and therefore the requisite information about the subject is available at the level of the sentence. Other examples include the problem of including anaphors in the complement position of certain prepositions in the binding domain of the head selecting the preposition (Pollard and Sag 1994) and recent proposals by Przepiórkowski (2000) and others to propagate the ARG-ST feature to phrases. In most cases, the authors cited provide some independent motivation for considering the information that is passed up a legitimate property of the embedded head. However, the hypothesis of locality of selection appears to be weakened to the extent that it is circumvented with purely formal (or diacritic) feature passing.

On the other hand, it is not clear that people do store statistical information about non-local subtrees. Gibson and Pearlmutter (1998) claim that people need only store frequentistic information about lexical items. According to Gibson and Pearlmutter, the strongest evidence to date for the need to keep track of frequentistic information about phrase structure is the different biases towards interpreting *that* as a complementizer or demonstrative article depending on its position in the phrase structure (33).

(33) a. That cheap hotel was clean and comfortable to our surprise.
b. That cheap hotels were clean and comfortable surprised us.
c. The lawyer insisted that cheap hotels were clean and comfortable.
d. The lawyer insisted that cheap hotel was clean and comfortable.

In sentence-initial position, *that* is more likely to be a demonstrative, making (33a) easier to process than (33b). After a verb, however, *that* is more likely to be a complementizer, making (33c) easier to process than (33d). Note that these biases involve non-local trees, as shown in Figure 6.5.

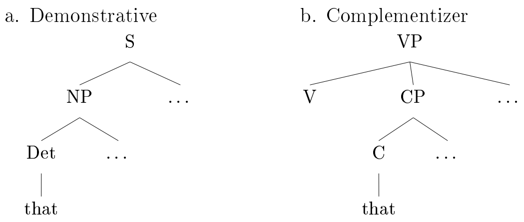


Figure 6.5: Trees encoding the dependency between interpretation of *that* and its syntactic context

However, Gibson and Pearlmutter (1998) suggest an alternative account, where memory resource factors (see Gibson 1998) would favor the article interpretation in sentence initial position. This is presumably because sentential subjects generally require the parser to posit more constituents before the main verb is encountered. The second part of their proposed account is that memory factors would not differentiate the two possibilities in the post-verbal context, and so straight lexical frequencies could apply, favoring the complementizer interpretation.

Further, it is not the case that allowing some frequency tabulation on non-local trees would necessarily violate locality irreparably. Frequency information about verbs and the prepositions they co-occur with, for example, would simply mirror one of the 'exceptions' to locality that is already countenanced. Note also that this proposal applies not only to frequency information about non-local trees, but also the suggestion of storing certain overspecified types in the grammar. For example, the phrase *I don't know*, which phonological effects suggest is a part of the grammar, spans a tree two levels deep. Likewise, idioms like *kick the bucket*, which must be in the grammar by the arbitrariness criterion, can also include multiple levels of phrase structure. The fact that some idioms are discontinuous (*put* x_i's *foot in* x_i's *mouth*) and the fact that idioms generally conform to the syntactic patterns of the language argue against giving them unusual flat structures. Thus it could be that

there is some principle to the storage of non-local trees that nonetheless preserves the core predictions of locality. Perhaps such trees must be fully specified, or if they do preserve only some information about lexical items, those items must be heads.[26] This would rule out trees encoding, say, a dependency between a verb and the case of an NP buried down inside its complement.

I would like to note in this regard that Bod's (1998) arguments for preserving all of the subtrees of previously encountered parses are based on the performance of a computational system. This system does a better job of disambiguation when it has access to all of the subtrees of the training corpus. Any systematic reduction in trees (including a reduction to trees lexified only with heads such as that mentioned above) hurts its performance. What is not clear from Bod's experiments is to what extent saving all of the trees serves to compensate for machines' lack of world knowledge and other contextual information. It is well known that humans rely on both for disambiguation and they therefore may need to store less than the computer (see, e.g., Pearlmutter and MacDonald 1992). Further, even if storing more would improve human systems, it doesn't follow that we do so.

The bulk of this section has been spent on exploring how the proposed test is complicated by circumstances, but I believe that this test or another (focusing on a different kind of constraint on grammars) could in principle be used to determine whether knowledge of language of the statistical sort must be kept separate from the (rest of) grammar.

6.6 Grammar as practice

In this concluding section, I would like to suggest that the model of grammar sketched in this chapter (and similarly, the usage-based model of Langacker 1987) is a promising candidate for a theory of knowledge of language within the social theory of language that Eckert (2000) calls for. After describing the connection that I see, I will end with a brief note about the status of grammaticality judgments on this view of grammar.

[26]This won't cover all cases of idioms, however (Susanne Riehemann, p.c.).

6.6.1 Linguistic knowledge in a social theory of language

Eckert articulates some of the requirements on a theory of grammar as practice:

> It is impossible for a social theory of language to view *langue* as a pre-existing convention, for a social theory of language must be about the process of conventionalization. By the same token, it is impossible for a social theory of language to view the individual speaker's competence as a simple internalization of convention. Convention and individual competence are mutually produced and reproduced in practice, thus linguistic practice is not simply the consensual use of a common system. Convention is not a thing but a process, and the possibility of convention resides in speakers' ability to hypothesize about others' behavior and to take interpretable action, along with a commitment to doing so within a particular social unit. Our speaker, or speaking subject, can not be a clone but must be an agent in a process of convention-making. (2000:45)

All of the suggestions for changes to the conception of grammar presented in this chapter can be seen as following from the view that convention is grounded in practice. If speakers' knowledge of language derives from (and is constantly fed by) their experience of situated language use, then the social aspects of those situations should be able to attach to particular signs. Further, the claim that convention is a process includes the idea that language is always changing. Each utterance potentially redefines or reinforces the system of the interlocutors attending to it. If Eckert and others are right (and they seem to be) that change in language structure is driven by the social actions that speakers carry out through their choice of forms, then the social meaning of linguistic forms must be represented somehow. Further, if individual utterances are what slowly reinforce or redefine the grammar, then the possibility arises that speakers store both overspecified types (e.g., common utterance types) and frequentistic information.

Note however, that this view of grammar is not meant to be in itself a social theory of language. Nor is it a theory of the process of conventionalization. Rather, it is a model of what linguistic knowledge should look like, given that "convention is not a thing but a process."

6.6.2 Problematizing grammaticality judgments

In shifting the definition of grammar away from concerns about grammaticality, it seems appropriate to reconsider grammaticality judgments as well. Langacker writes:

> The notion of syntax as an autonomous formal system has encouraged the expectation that speakers should be capable of simple categorical judgments (grammatical/ungrammatical) on the well-formedness of sentences, out of context and without regard for semantic considerations: either a sentence meets all formal specifications, or it does not. (1987:36)

Grammaticality judgments have been conceptualized/naturalized as a property of the brain, partially hidden from observation by performance factors. Linguists and, to some extent, people we train as consultants share a set of beliefs that allow us to ask and answer these questions. It does not follow that there is some underlying system which generates judgments that are then obscured. To the extent that the effect of asking for judgments on decontextualized examples has been acknowledged as a problem, it has generally been considered as simply a confounding factor, obscuring the otherwise pristine data of grammaticality judgments. From a grammar as practice viewpoint, however, the fact that speakers need to be given or generate some context in order to make sense of examples submitted for judgment indicates that the activity of giving grammaticality judgments is derivative. The most natural use of our linguistic systems is in situated conversation.

From a grammar as practice viewpoint, one could say that speakers have beliefs like "we say it thus when we mean this," based on experience. These facts are naturalized for speakers in the sense that they come to believe that it is an essential property of speech (or, if they reflect on it, of their language), that things are said in a certain way. This process of naturalization gives the beliefs some inertia or persistence. However, enough experience (especially *thrown* interaction involving language) can provide new models. This is not to deny patterns in language, but the judgment that something never said is out or bad is a different matter. Grammaticality judgments reflect the linguistic system, but are not the essence of it.

6.7 Conclusion

In this chapter, I have considered the boundaries of linguistic competence in light of the results presented in Chapters 4 and 5. My goal in this chapter has not been to definitively establish where exactly the boundaries of competence lie. Rather, I wanted to raise some possibilities not usually considered in generative theory and therefore highlight the fact that the question remains unsettled. The mainstream position, that competence is merely a matter of grammaticality, has been accepted but not proven, and these issues merit further investigation. I hope that this dissertation has shown that sociolinguistic variation, the social value of variables and the non-categorical grammatical constraints that apply to them provide an interesting locus for the study of the boundaries of linguistic competence.

Bibliography

Ades, A., and M.J. Steedman. 1982. On the order of words. *Linguistics and Philosophy* 4:517–558.

Akmajian, Adrian, Susan M. Steele, and Thomas Wasow. 1979. The category AUX in universal grammar. *Linguistic Inquiry* 10(1):1–64.

Andrews, Avery. 1982. The representation of case in modern Icelandic. In J. Bresnan (Ed.), *The Mental Representation of Grammatical Relations*, 427–503. Cambridge, MA: MIT Press.

Andrews, Avery. 1990. Case structures and control in modern Icelandic. In J. Maling and A. Zaenen (Eds.), *Modern Icelandic Syntax*, 187–234. New York: Academic Press.

Anisfeld, E., and Wallace E. Lambert. 1964. Evaluational reactions of bilingual and monolingual children to spoken language. *Journal of Abnormal and Social Psychology* 69:89–97.

Anttila, Arto. 1997. Deriving variation from grammar. In F. Hinskens, R. van Hout, and W. L. Wetzels (Eds.), *Variation, Change and Phonological Theory*, 35–68. Amsterdam: John Benjamins.

Bailey, Guy. 1987. Contribution to panel discussion on "are black and white vernaculars diverging?". *American Speech* 62:32–40.

Bailey, Guy, and Natalie Maynor. 1985. The present tense of *be* in White folk speech of the Southern United States. *English World-Wide* 6:199–216.

Bailey, Guy, and Natalie Maynor. 1987. Decreolization? *Language in Society* 16:449–473.

Bakhtin, Mikhail Mikhaĭlovich. 1986. *Speech Genres and Other Late Essays*. Austin: University of Texas Press. Translated by Vern W. McGee. Edited by Caryl Emerson and Michael Holquist.

Balari, Sergio. 1992. Agreement and θ-roles: Towards and account of null subjects in hpsg. CLAUS report 25, Computerlinguistik an der Universität des Saarlandes, Saarbrücken.

Baugh, John. 1979. *Linguistic Style Shifting in Black English*. PhD thesis, University of Pennsylvania.

Bender, Emily. 1998. Are there any side effects to that prescription? Unpublished ms., Paper presented at NWAV(E) 27, University of Georgia, Athens.

Bender, Emily. 1999. Constituting context: Null objects in English recipes revisited. In J. Alexander, N.-R. Han, and M. M. Fox (Eds.), *Penn Working Papers in Linguistics* 6(1), 53–68.

Bender, Emily, and Dan Flickinger. 1999. Peripheral constructions and core phenomena: Agreement in tag questions. In G. Webelhuth, J.-P. Koenig, and A. Kathol (Eds.), *Lexical and Constructional Aspects of Linguistic Explanation*, 199–214. Stanford, CA: CSLI.

Bickerton, Derek. 1971. Inherent variability and variable rules. *Foundations of Language* 7:457–492.

Black, Ezra, Fred Jelinek, John Lafferty, David M. Magerman, Robert Mercer, and Salim Roukos. 1993. Towards history-based grammars: Using richer models for probabilistic parsing. In *ACL 31*, 31–37.

Bloch, B. 1948. A set of postulates for phonemic analysis. *Language* 24:3–46.

Bod, Rens. 1998. *Beyond Grammar: An Experience-Based Theory of Language.* Stanford, CA: CSLI.

Boersma, Paul, and Bruce Hayes. 1999. Empirical tests of the gradual learning algorithm. Unpublished ms., University of Amsterdam and UCLA.

Bouma, Gosse, Rob Malouf, and Ivan A. Sag. in press. Satisfying constraints on extraction and adjunction. *Natural Language and Linguistic Theory.*

Brauner, Siegmund. 1974. *Lehrbuch des Bambara.* Leipzig: VEB Verlag Enzyklopädie.

Bresnan, Joan. 1978. A realistic transformational grammar. In M. Halle, J. Bresnan, and G. Miller (Eds.), *Linguistic Theory and Psychological Reality*, 1–59. Cambridge, MA: MIT Press.

Bresnan, Joan. 1982. Control and complementation. *Linguistic Inquiry* 13(3):343–434.

Bresnan, Joan, and Sam A. Mchombo. 1987. Topic, pronoun, and agreement in Chichewa. In M. Iida, S. Wechsler, and D. Zec (Eds.), *Working Papers in Grammatical Theory and Discourse Structure: Interactions of Morphology, Syntax and Discourse*, 1–59. Stanford, CA: CSLI.

Brown, Penelope, and Stephen C. Levinson. 1987. *Politeness: Some Universals in Language Usage.* Cambridge University Press: Cambridge.

Butters, Ronald R. 1972. Competence, performance, and variable rules. *Language Sciences* 20:29–31.

Bybee, Joan, and Joanne Scheibman. 1999. The effect of usage on degrees of constituency: The reduction of *don't* in English. *Linguistics* 37(4):575–596.

California Style Collective. 1993. Variation and personal/group style. Unpublished ms., Paper presented at NWAVE-22.

Cedergren, Henrietta J. 1973. *Interplay of Social and Linguistic Factors in Panama.* PhD thesis, Cornell University.

Cedergren, Henrietta J., and David Sankoff. 1974. Variable rules: Performance as a statistical reflection of competence. *Language* 50(2):333–355.

Chambers, J. K. 1995. *Sociolinguistic Theory: Linguistic Variation and its Social Significance.* Oxford: Blackwell.

Charniak, Eugene. 1996. *Statistical Language Learning.* Cambridge, MA: MIT Press.

Chomsky, Noam. 1955. The logical structure of linguistic theory. Unpublished ms., Published 1975 by New York: Plenum.

Chomsky, Noam. 1981. *Lectures on Government and Binding.* Dordrecht: Foris.

Chomsky, Noam. 1986. *Knowledge of Language: Its Nature, Origin, and Use.* New York: Praeger.

Chomsky, Noam. 1991. Some notes on economy of derivation and representation. In R. Friedin (Ed.), *Principles and Parameters in Comparative Grammar*, 417–454. Cambridge, MA: MIT Press.

Chomsky, Noam, and Morris Halle. 1968. *The Sound Pattern of English.* New York: Harper & Row.

Clark, Herbert H. 1996. *Using Language.* Cambridge: Cambridge University Press.

Cohen, Jonathan D., Brian MacWhinney, Matthew Flatt, and Jefferson Provost. 1993. PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments, and Computers* 25:257–271.

Collins, Michael John. 1996. A new statistical parser based on bigram lexical dependencies. In *ACL 34*, 185–191.

Collins, Peter. 1992. Cleft existentials in English. *Language Sciences* 14(4):419–433.

Connell, R. W. 1995. *Masculinities.* Berkeley: University of California Press.

Copestake, Anne, Dan Flickinger, Ivan A. Sag, and Carl Pollard. 1999. Minimal recursion semantics: An introduction. Unpublished ms., Stanford University and OSU.

Coupland, Nikolas. in press. Language, situation and the relational self: Theorising dialect-style in sociolinguistics. In R. John and P. Eckert (Eds.), *Stylistic Variation*. Cambridge: Cambridge University Press.

Craig, Collette Grinevald. 1977. *The Structure of Jacaltec*. Austin: University of Texas Press.

Crain, Stephen, and Mark Steedman. 1985. On not being led up the garden path: The use of context by the psychological syntax processor. In D. R. Dowty, L. Karttunen, and A. M. Zwicky (Eds.), *Natural Language Parsing: Psychological, Computational, and Theoretical Perspectives*, 320–358. Cambridge: Cambridge University Press.

Davis, Anthony. 1996. *Lexical Semantics and Linking in the Hierarchical Lexicon*. PhD thesis, Stanford University.

Davis, Anthony R., and Jean-Pierre Koenig. 2000. Linking as constraints on world classes in a hierarchical lexicon. *Language* 76(1):56–91.

Day, Richard R. 1973. *Patterns of Variation in Copula and Tense in the Hawaiian Post-Creole Continuum*. Vol. 5.2 of *Working Papers in Linguistics*. Department of Linguistics, University of Hawaii.

Debose, Charles. 1994. A note on *ain't* vs. *didn't* negation in African American Vernacular. *Journal of Pidgin and Creole Languages* 9:127–130.

Dittmar, Norbert. 1996. Descriptive and explanatory power of rules in sociolinguistics. In R. Singh (Ed.), *Towards a Critical Sociolinguistics*, 115–149. Philadelphia: John Benjamins.

Eckert, Penelope. 2000. *Linguistic Variation as Social Practice*. Oxford: Blackwell.

Eckert, Penelope, and Sally McConnell-Ginet. 1992. Think Practically and Look Locally: Language and gender as community-based practice. *Annual Review of Anthropology*.

Epstein, Julia. 1990. Either/or—neither/both: Sexual ambiguity and the ideology of gender. *Genders* 7:99–142.

Falk, Yehuda N. 1984. The English auxiliary system: A Lexical-Functional analysis. *Language* 60(3):483–509.

Fasold, Ralph W. 1991. The quiet demise of variable rules. *American Speech* 66(1):3–21.

Fasold, Ralph W. 1996. The quiet demise of variable rules. In R. Singh (Ed.), *Towards a Critical Sociolinguistics*, 79–97. Philadelphia: John Benjamins.

Ferguson, Charles A. 1971. Absence of copula and the notion of simplicity: A study of normal speech, baby talk, foreigner talk, and pidgins. In D. Hymes (Ed.), *Pidginization and Creolization of Languages*, 141–150. New York: Cambridge University Press.

Ferguson, Charles A. 1972. 'being' in Bengali, with a note on Amharic. In J. M. W. Verhaar (Ed.), *The Verb "Be" and Its Synonyms*, 74–114. Dordrecht: Reidel.

Fillmore, Charles J. 1987. Varieties of conditional sentences. In F. Marshall, A. Miller, and Z.-s. Zhang (Eds.), *Proceedings of the Third Eastern States Conference on Linguistics*, 163–182. Pittsburgh, PA: The University of Pittsburgh.

Fillmore, Charles J. 1988. The mechanisms of "Construction Grammar". In S. Axmaker, A. Jaisser, and H. Singmaster (Eds.), *Berkeley Linguistics Society Proceedings of the Fourteenth Annual Meeting*, 35–56. Berkeley, CA: Berkeley Linguistics Society.

Fillmore, Charles J. 1999. Inversion and constructional inheritance. In G. Webelhuth, J.-P. Koenig, and A. Kathol (Eds.), *Lexical and Constructional Aspects of Linguistic Explanation*, 113–128. Stanford, CA: CSLI.

Fillmore, Charles J., Paul Kay, and Mary Catherine O'Connor. 1988. Regularity and idiomaticity in grammatical constructions: The case of let alone. *Language* 64(3):501–538.

Fischer, John L. 1958. Social influences on the choice of a linguistic variant. *Word* 14:47–56.

Flickinger, Danieal, Carl Pollard, and Thomas Wasow. 1985. Structure sharing in lexical representations. In *Proceedings of the Twenty-third Annual Meeting of the ACL*, 262–267, Morristown, NJ. Association for Computational Linguistics.

Flickinger, Daniel. 1987. *Lexical Rules in the Hierarchical Lexicon.* PhD thesis, Stanford University.

Fries, Charles C., and Kenneth L. Pike. 1949. Coexistent phonemic systems. *Language* 25.

Gazdar, Gerald. 1976. Quantifying context. *York Papers in Linguistics* 6:117–129.

Gazdar, Gerald. 1981. Unbounded dependencies and coordinate structure. *Linguistic Inquiry* 12(2):155–184.

Gazdar, Gerald, Ewan Klein, Geoffrey Pullum, and Ivan Sag. 1985. *Generalized Phrase Structure Grammar*. Cambridge, MA: Harvard Univrsity Press.

Gazdar, Gerald, Ewan Klein, Geoffrey Pullum, and Ivan A. Sag. 1984. Foot features and parasitic gaps. In W. d. Geest and Y. Putseys (Eds.), *Sentential Complementation*, 83–94. Dordrecht: Foris.

Gazdar, Gerald, Geoffrey Pullum, and Ivan A. Sag. 1982. Auxiliaries and related phenomena in a restricted theory of grammar. *Language* 58:591–638.

Gibson, Edward. 1998. Linguistic complexity: Locality of syntactic dependencies. *Cognition* 68:1–76.

Gibson, Edward, and Neal J. Pearlmutter. 1998. Constraints on sentence comprehension. *Trends in Cognitive Sciences* 2(7):262–268.

Giles, Howard. 1971. Our reactions to accents. *New Society* (October 14th):713–715.

Ginzburg, Jonathan, and Ivan A. Sag. to appear. *English Interrogative Constructions*. Stanford, CA: CSLI. Hebrew University of Jerusalem and Stanford University.

Grant, D. A., H. W. Hake, and J. P. Hornseth. 1951. Acquisition and extinction of a verbal conditioned response with differing percentages of reinforcement. *Journal of Experimental Psychology* 52:1–5.

Green, Lisa. in preparation. Patterns in African American English: Description, practice and problems. Unpublished ms.

Guy, Gregory R. 1980. Variation in the group and in the individual: The case of final stop deletion. In W. Labov (Ed.), *Locating Language in Time and Space*. New York: Academic Press.

Guy, Gregory R. 1988. Advanced Varbrul analysis. In K. Ferrara, B. Brown, K. Walters, and J. Baugh (Eds.), *Linguistic Change and Contact*, 124–136. Austin, Texas: University of Texas Department of Linguistics.

Guy, Gregory R. 1996. Form and function in linguistic variation. In G. R. Guy, C. Feagin, D. Schiffrin, and J. Baugh (Eds.), *Towards a Social Science of Language: Papers in Honor of William Labov*, Vol. 1: *Variation and Change in Language and Society*, 221–252. Amsterdam: John Benjamins.

Halliday, M. A. K. 1991. Corpus studies and probabilistic grammar. In K. Aijmer and B. Altenberg (Eds.), *English Corpus Linguistics: Studies in Honour of Jan Svartvik*, 30–43. Longman.

Hannah, Dawn. 1996. Copula absence in Samaná English. Unpublished ms., PhD qualifying paper, Department of Linguistics, Stanford University. Revised version to apppear in *American Speech*.

Harder, Peter. 1996. *Functional Semantics: A Theory of Meaning, Structure and Tense in English*. Berlin: Mouton de Gruyter.

Harder, Peter. 2000. The status of linguistic facts: Rethinking the relation between cognition, social institution, and utterance from a functional point of view. Unpublished ms., University of Copenhagen.

Haser, I., and W. Chromiak. 1977. The processing of frequency: An automatic mechanism? *Journal of Verbal Learning and Verbal Behavior* 16:173–184.

Henderson, L. 1989. On mental representation of morphology and its diagnosis by measures of visual access speed. In W. Marslen-Wilson (Ed.), *Lexical Representations and Processes*, 357–391. Cambridge, MA: MIT Press.

Hendrick, Randall. 1982. Reduced questions and their theoretical implications. *Language* 58:800–819.

Hendry, Joy. 1992. Honorifics as dialect: The expression and manipulation of boundaries in japanese. *Multilingua* 11(4):341–354.

Hengeveld, Kees. 1992. *Non-verbal Predication: Theory, Typology, Diachrony.* New York: Mouton de Gruyter.

Horn, Laurence R. 1972. *On the Semantic Properties of Logical Operators in English.* PhD thesis, UCLA. Distributed by IULC, 1976.

Hudson, Richard A. 1996. *Sociolinguistics*. Cambridge: Cambridge University Press. Second edition.

Irvine, Judith T. 1990. Registering affect: Heteroglossia in the linguistic expression of emotion. In C. A. Lutz and L. Abu-Lughod (Eds.), *Language and the Politics of Emotion*. Cambridge: Cambridge University Press.

Irvine, Judith T. in press. "style" as distinctiveness: The culture and ideology of linguistic differentiation. In R. John and P. Eckert (Eds.), *Stylistic Variation*. Cambridge: Cambridge University Press.

Irvine, Judith T., and Susan Gal. 2000. Language ideology and linguistic differentiation. In P. V. Kroskrity (Ed.), *Regimes of Language: Ideologies, Polities, and Identities*, 35–83. Santa Fe, NM: SAR Press.

Jaeggli, Osvaldo. 1982. *Topics in Romance Syntax*. Dordrecht: Foris.

Johnson, Mark, and Martin Kay. 1994. Parsing and empty nodes. *Computational Linguistics* 20(2):289–299.

Jurafsky, Daniel. 1996. A probabilistic model of lexical and syntactic access and disambiguation. *Cognitive Science* 20:137–194.

Kaplan, Ronald M., and Annie Zaenen. 1988. Long-distance dependencies as a case of functional uncertainty. In M. R. Baltin and A. S. Kroch (Eds.), *Alternative Conceptions of Phrase Structure Grammar*, 17–42. Chicago: University of Chicago Press.

Kathol, Andreas. 1999. Agreement and the syntax-morphology interface in HPSG. In R. Levine and G. Green (Eds.), *Studies in Current Phrase Structure Grammar*. Cambridge University Press.

Kay, Paul, and Chad K. McDaniel. 1979. On the logic of variable rules. *Language in Society* 8(2):151–187.

Kemmer, Suzanne, and Michael Barlow. 2000. Introcution: A usage-based conception of language. In M. Barlow and S. Kemmer (Eds.), *Usage-Based Models of Language*, vii–xxviii. Stanford: CSLI.

Kemmer, Suzanne, and Michael Israel. 1994. Variation and the usage-based model. *CLS* 30:165–179.

Kim, Jong-Bok. 2000. *The Grammar of Negation: A Constraint-Based Approach*. Stanford, CA: CSLI.

Kim, Jong-Bok, and Ivan A. Sag. 1995. The parametric variation of French and English negation. In J. Camacho, L. Choueiri, and M. Watanabe (Eds.), *Proceedings of the Fourteenth West Coast Conference on Formal Linguistics (WCCFL)*, 303–317. Stanford, CA: SLA CSLI Publications.

Kiparsky, Paul. 1972. Explanation in phonology. In S. Peters (Ed.), *Goals of Linguistic Theory*. Englewood Cliffs NJ: Prentice-Hall, Inc.

Kiparsky, Paul. 1988. Phonological change. In J. Newmeyer, Frederick (Ed.), *Linguistics: The Cambridge Survey*, Vol. Vol. 1: Linguistic Theory: Foundations, 363–415. Cambridge: Cambridge University Press.

Klima, E. 1964. Negation in English. In J. Fodor and J. Katz (Eds.), *The Structure of Language*. Englewood Cliffs: Prentice Hall.

Koenig, Jean-Pierre. 1999. *Lexical Relations*. Stanford, CA: CSLI.

Kortenhoven, Andrea. 2000. Raising their voices: Black girls' verbal culture in the face of hegemonic femininity. Unpublished ms., Stanford University.

Labov, William. 1963. The social motivation of a sound change. *Word* 19:273–309.

Labov, William. 1966. *The Social Stratification of English in New York City*. Washington, DC: Center for Applied Linguistics.

Labov, William. 1969. Contraction, deletion, and inherent variability of the English copula. *Language* 45(4).

Labov, William. 1972. *Language in the Inner City: Studies in the Black English Vernacular*. Philadelphia: University of Pennsylvania Press.

Labov, William. 1995. The case of the missing copula: The interpretation of zeroes in African-American English. In L. R. Gleitman and M. Liberman (Eds.), *An Invitation to Cognitive Science, Volume 1: Language*, 25–54. Cambridge, MA: MIT, 2nd edition.

Labov, William, Paul Cohen, Clarence Robins, and John Lewis. 1968. A study of the non-standard English of Negro and Puerto Rican speakers in New York City. Final Report, Cooperative Research Project No. 3288, United States Office of Education.

Lambert, Sylvie. 1973. The role of speech in forming evaluations: A study of children and teachers. Master's thesis, Tufts University.

Lambert, Wallace E. 1967. A social psychology of bilingualism. *Journal of Social Issues* 23:91–109.

Lambert, Wallace E., H. Frankel, and G. R. Tucker. 1966. Judging personality through speech: A French-Canadian example. *Journal of Communication* 16:305–321.

Lambert, Wallace E., Howard Giles, and Omer Picard. 1975. Language attitudes in a French-American community. *Linguistics* 158:127–152.

Lambert, Wallace E., and G. Richard Tucker. 1975. White and Negro listeners' reactions to various American-English dialects. In J. Dillard (Ed.), *Perspectives on Black English*, Vol. 4 of *Contributions to the Sociology of Language*, 369–377. The Hague: Mouton.

Langacker, Ronald W. 1987. *Foundations of Cognitive Grammar, Vol I: Theoretical Prerequisites*. Stanford, CA: Stanford University Press.

Langacker, Ronald W. 1990. *Concept, Image, and Symbol: The Cognitive Basis of Grammar*. New York: Mouton de Gruyter.

Langacker, Ronald W. 2000. A dynamic usage-based model. In M. Barlow and S. Kemmer (Eds.), *Usage-Based Models of Language*, 1–63. Stanford: CSLI.

Lapointe, Steven G. 1980. A lexical analysis of the English auxiliary verb system. In T. Hoekstra, van der Hulst H., and M. Moortgat (Eds.), *Lexical Grammar*, 215–254. Dordrecht: Foris.

Lasnik, Howard. 1992. Traces. In W. Bright (Ed.), *International Encyclopedia of Linguistics*, Vol. 4. Oxford: Oxford University Press.

Lavandera, Beatriz R. 1978. Where does the sociolinguistic variable stop? *Language in Society* 7:171–182.

Levin, Nancy S. 1986. *Main Verb Ellipsis in Spoken English*. New York: Garland.

Lewis, David K. 1969. *Convention: A Philosophical Study*. Cambridge, MA: Harvard University Press.

Lumsden, Michael. 1988. *Existential Sentences: Their Structure and Meaning*. London: Croom Helm.

MacDonald, Maryellen C. 1994. Probabilistic constraints and syntactic ambiguity resolution. *Language and Cognitive Processes* 9:157–201.

MacDonald, Maryellen C., Neal J. Pearlmutter, and Mark S. Seidenberg. 1994. Lexical nature of syntactic ambiguity resolution. *Psychological Review* 101(4):676–703.

Maddieson, Ian. 1984. The effects on F0 of a voicing distinction in sonorants and their implications for a theory of tonogenesis. *Journal of Phonetics* 12(1):9–15.

Manning, Christopher D. 1996. *Ergativity: Argument Structure and Grammatical Relations*. Stanford, CA: CSLI.

Manning, Christopher D., and Ivan A. Sag. 1998. Argument structure, valence, and binding. *Nordic Journal of Linguistics* 21(2):107–144.

Manning, Christopher D., and Ivan A. Sag. 1999. Dissociations between argument structure and grammatical relations. In G. Webelhuth, J.-P. Koenig, and A. Kathol (Eds.), *Lexical and Constructional Aspects of Linguistic Explanation*, 63–78. Stanford, CA: CSLI.

Manning, Christopher D., and Hinrich Schütze. 1999. *Foundations of Statistical Natural Language Processing*. Cambridge, MA: MIT Press.

McCawley, James D. 1988. The comparative conditional construction in English, German, and Chinese. In S. Axmaker, A. Jaisser, and H. Singmaster (Eds.), *Berkeley Linguistics Society Proceedings of the Fourteenth Annual Meeting*, 176–187. Berkeley, CA: Berkeley Linguistics Society.

McConnell-Ginet, Sally. 1984. The origins of sexist language in discourse. In S. J. White and V. Teller (Eds.), *Discourses in Reading and Linguistics*, 123–135. New York: New York Academy of Sciences.

McConnell-Ginet, Sally. 1989. The sexual (re)production of meaning: A discourse-based theory. In F. W. Frank and P. A. Treichler (Eds.), *Language, Gender, and Professional Writing*, 35–50. MLA.

Miller, Philip. 1992. *Clitics and Constituents in Phrase Structure Grammar*. New York: Garland.

Milsark, Gary L. 1974. *Existential Sentences in English*. PhD thesis, MIT. Published by Garland, 1979.

Morgan, J. L. 1978. Two types of convention in indirect speech acts. *Syntax and Semantics* 9:261–280.

Mufwene, Salikoko S. 1992. Why grammars are not monolithic. In D. Brentari, G. N. Larson, and L. A. MacLeod (Eds.), *The Joy of Grammar*, 225–250. Amsterdam: John Benjamins.

Neumann, Günter, and Dan Flickinger. 1999. Learning stochastic lexicalized tree grammars from HPSG. Technical report, DFKI.

Nordlinger, Rachel. 1998. *Constructive Case*. Stanford, CA: CSLI.

Ochs, Elinor. 1992. Indexing gender. In A. Duranti and C. Goodwin (Eds.), *Rethinking Context: Language as an Interactive Phenomenon*, 335–358. Cambridge: Cambridge University Press.

Okushi, Yoshiko. 1997. *Patterns of Honorific Use in the Everyday Speech of Four Japanese Women*. PhD thesis, University of Pennsylvania.

Pearlmutter, N., and M. MacDonald. 1992. Plausibility and syntactic ambiguity resolution. In *Proceedings 14th Annual Conference of the Cognitive Society*.

Pickering, Martn, and Guy Barry. 1991. Sentence processing without empty categories. *Language and Cognitive Processes* 6(3):229–259.

Pollard, Carl. 1989. The syntax-semantics interface in a unification-based phrase structure grammar. In S. Busemann, C. Hauenschild, and C. Umbach (Eds.), *Views of the Syntax/Semantics Interface*. Berlin: Technical University of Berlin. No. 74 KIT-FAST Report.

Pollard, Carl, and Ivan A. Sag. 1987. *Information-Based Syntax and Semantics, Vol I – Fundamentals*. Stanford, CA: CSLI.

Pollard, Carl, and Ivan A. Sag. 1994. *Head-driven Phrase Structure Grammar*. Chicago: Chicago University Press.

Pollard, Carl J. 1984. *GPSGs, Head Grammar, and Natural Language*. PhD thesis, Stanford.

Poplack, Shana, and David Sankoff. 1987. The Philadelphia story in the Spanish Caribbean. *American Speech* 62:291–314.

Poplack, Shana, and Sali Tagliamonte. 1991. African American English in the diaspora: Evidence from old-line Nova Scotians. *Language Variation and Change* 3:301–339.

Preston, Dennis R. 1989. *Sociolinguistics and Second Language Acquisition*. Oxford: Blackwell.

Prince, Alan. 2000. The special and the general. Unpublished ms., Paper presented at the 24th Penn Linguistics Colloquium.

Prince, Alan, and Paul Smolensky. 1993. Optimality Theory: Constraint interaction in generative grammar. Unpublished ms., Rutgers University and University of Colorado, Boulder.

Przepiórkowski, Adam. 2000. ARG-ST on phrases headed by semantically vacuous words: Evidence from polish. Unpublished ms., Paper presented at the 7th International Conference on Head-Driven Phrase Structure Grammar.

Pullum, Geoffrey K., and Arnold M. Zwicky. 1997. Licensing of features by syntactic rules: The key to Auxiliary Reduction. Unpublished ms., Paper presented at the annual meeting of the LSA.

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech, and Jan Svartvik. 1985. *A Comprehensive Grammar of the English Language*. London: Longman.

Reynolds, Bill, and Naomi Nagy. 1994. Phonological variation in Faetar: An Optimality account. In K. Beals, J. Denton, R. Knippen, L. Melnar, H. Suzuki, and E. Zeinfeld (Eds.), *Papers from the 30th Regional Meeting of the Chicago Linguistic Society, Volume 2: The Parasession on Variation in Linguistic Theory*, 277–292.

Reynolds, W. T. 1994. *Variation and Phonological Theory*. PhD thesis, University of Pennsylvania.

Rickford, John R. 1992. The creole residue in Barbados. In J. H. Hall, N. Doane, and D. Ringler (Eds.), *Old English and New: Studies in Honor of Frederic G. Cassidy*, 183–201. New York: Garland.

Rickford, John R. 1996a. Copula variability in Jamaican Creole and African American Vernacular English: A reanalysis of DeCamp's texts. In G. R. Guy, C. Feagin, D. Schiffrin, and J. Baugh (Eds.), *Towards a Social Science of Language: Papers in Honor of William Labov*, Vol. 1: *Variation and Change in Language and Society*, 357–372. Amsterdam: John Benjamins.

Rickford, John R. 1996b. Variation in the Jamaican Creole copula: New data and analysis. Unpublished ms., Paper presented at the Beryl Bailey Symposium, American Anthropology Association meeting, Chicago.

Rickford, John R. 1998. The creole origins of African-American vernacular English: Evidence from copula absence. In S. S. Mufwene, J. R. Rickford, G. Bailey, and J. Baugh (Eds.), *African-American English*, 154–200. New York: Routledge.

Rickford, John R., Arnetha Ball, Renee Blake, Raina Jackson, and Nomi Martin. 1991. Rappin on the copula coffin: Theoretical and methodological issues in the analysis of copula variation in African-American Vernacular English. *Language Variation and Change* 3:103–132.

Rickford, John R., and Renee Blake. 1990. Copula contraction and absence in Barbadian English, Samanà English and Vernacular Black English. In K. Hall, J.-P. Koenig, M. Meacham, S. Reinman, and L. A. Sutton (Eds.), *Proceedings of the Sixteenth Annual Meeting of the Berkeley Linguistics Society*, 257–268.

Rickford, John R., and Faye McNair-Knox. 1994. Addressee- and topic-influenced style shift: A quantitative sociolinguistic study. In D. Biber and E. Finegan (Eds.), *Sociolinguistic Perspectives on Register*, 235–276. New York: Oxford University Press.

Riehemann, Susanne. forthcoming. *A Constructional Approach to Idioms and Word Formation.* PhD thesis, Stanford University.

Riehemann, Susanne Z. 1999. Type-based derivational morphology. *Journal of Comparative Germanic Linguistics* 2:49–77.

Rizzi, Luigi. 1982. *Issues in Italian Syntax*. Dordrecht: Foris.

Romaine, Suzanne. 1981. *On the Problem of Syntactic Variation: A Reply to Beatriz Lavandera and William Labov.* Austin, Texas: Southwest Educational Development Laboratory.

Romaine, Suzanne. 1982. *Socio-Historical Linguistics: Its Status and Methodology.* Cambridge: Cambridge University Press.

Rosenbaum, Peter S. 1967. *The Grammar of English Predicate Complement Constructions*. Cambridge, MA: MIT Press.

Ross, John R. 1967a. Axuiliaries as main verbs. *Studies in Philosophical Linguistics* Series 1.

Ross, John R. 1967b. *Constraints on Variables in Syntax.* PhD thesis, MIT. (Published as *Infinite Syntax!*, Norwood, NJ: Ablex, 1986).

Rousseau, Pascale, and David Sankoff. 1978. Advances in variable rule methodology. In D. Sankoff (Ed.), *Linguistic Variation: Models and Methods*, 57–69. New York: Academic Press, Inc.

Sag, Ivan A. 1992. Taking performance seriously. In C. Martin-Vide (Ed.), *VII Congresso de Languajes Naturales y Lenguajes Formales*.

Sag, Ivan A. 1997. English relative clause constructions. *Journal of Linguistics* 33(2):431–483.

Sag, Ivan A. 1998. Without a trace. Unpublished ms., Stanford University.

Sag, Ivan A. 1999. Explaining the English auxiliary system. Unpublished ms., Paper presented at the UCSD Seminar on the Nature of Explanation in Linguistic Theory.

Sag, Ivan A., and J. Fodor. 1994. Extraction without traces. In *Proceedings of WCCFL 13*, 122–138. Stanford, CA: CSLI.

Sag, Ivan A., and Thomas Wasow. 1999. *Synactic Theory: A Formal Introduction.* Stanford, CA: CSLI.

Sankoff, David. 1975. Varbrul version 2. Unpublished ms.

Schütze, Carson T. 1996. *The Empirical Base of Linguistics: Grammaticality Judgments and Linguistic Methodology.* Chicago: University of Chicago Press.

Searle, John R. 1995. *The Construction of Social Reality.* New York: The Free Press.

Sells, Peter, John Rickford, and Thomas Wasow. 1996. An optimality theoretic approach to variation in negative inversion in AAVE. *Natural Language and Linguistic Theory* 16(3):591–627.

Seuren, Pieter A. M. 1982. Internal variability in competence. *Linguistische Berichte* 77:1–31.

Singler, John. 1991. Copula variation in Liberian Settler English and American Black English. In W. F. Edwards and D. Winford (Eds.), *Verb Phrase Patterns in Black English and Creoles*, 129–164. Detroit: Wayne State University Press.

Soames, Scott. 1984. Linguistics and psychology. *Linguistics and Philosophy* 7:155–179.

Sylva, Clare M., and Arnold Zwicky. 1975. Discord. In R. W. Fasold and R. W. Shuy (Eds.), *Analyzing Variation in Language: Papers from the Second Colloquium on New Ways of Analyzing Variation*, 203–219. Washington, DC: Georgetown University Press.

Warner, Anthony. 1993. *English Auxiliaries: Structure and History.* Cambridge: Cambridge University Press.

Warner, Anthony. 2000. English auxiliaries without lexical rules. In R. Borsley (Ed.), *Syntax and Semantics 32: The Nature and Function of Syntactic Categories*, 167–220. New York: Academic Press.

Wasow, Thomas. 1972. *Anaphoric Relations in English.* PhD thesis, MIT.

Wasow, Thomas. 1978. On constraining the class of transformational languages. *Synthese* 39.

Weiner, E. Judith, and William Labov. 1983. Constraints on the agentless passive. *Journal of Linguistics* 19:29–58.

Weinreich, Uriel, William Labov, and Martin I. Herzog. 1968. Empirical foundations for a theory of language change. In W. P. Lehmann and Y. Malkiel (Eds.), *Directions for Historical Linguistics: A Symposium*, 95–188. Austin: University of Texas Press.

Weldon, Tracey. 1994. Variability in negation in African American Vernacular English. *Language Variation and Change* 6:359–397.

Williams, Edwin. 1984. *There*-insertion. *Linguistic Inquiry* 15(1):131–153.

Winford, Donald. 1992. Another look at the copula in Black English and Caribbean creoles. *American Speech* 67(1):21–60.

Winograd, Terry, and Fernando Flores. 1986. *Understanding Computers and Cognition: A New Foundation for Design.* New York: Addison-Wesley.

Wolfram, Walt. 1974. The relationship of White Southern Speech to Vernacular Black English. *Language* 50(3):498–527.

Wolfram, Walt. 1991. The linguistic variable: Fact and fantasy. *American Speech* 66(1):22–32.

Wolfram, Walter A. 1969. *A Sociolinguistic Description of Detroit Negro Speech.* Center for Applied Linguistics.

Wright, Saundra, and Jennifer Hay. 2000. Fred and wilma: A phonological conspiracy. Unpublished ms., Paper presented at IGALA 1.

Zwicky, Arnold M. 1982. Stranded *to* and phonological phrasing in english. *Linguistics* 20(1–2):3–57.

Zwicky, Arnold M. 1994. Dealing out meaning: Fundamentals of syntactic constructions. In S. Gahl, A. Dolbey, and C. Johnson (Eds.), *Proceedings of the Twentieth Annual Meeting of the Berkeley Linguistics Society*, 611–625.

Zwicky, Arnold M., and Geoffrey K. Pullum. 1983. Cliticization vs. inflection: English *n't*. *Language* 59:502–13.